

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2002, 20:40:02 ; Search time 51 Seconds
(without alignments)
772.845 Million cell updates/sec

Title: US-09-699-652A-14

Perfect score: 2194

Sequence: 1 MAMAGHAPGGALPLILLVWS.....DDVYVDLIRFLRENGWHNSY 410

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1055	48.1	344	2 E84526	probable lysosomal
2	557.5	25.4	403	2 T33198	hypothetical prote
3	553	25.2	397	1 JC4017	triacylglycerol li
4	549	25.0	405	2 T22675	hypothetical prote
5	546	24.9	398	2 S07145	triacylglycerol li
6	539.5	24.6	395	1 L1RTT	triacylglycerol li
7	523	23.8	405	2 H88930	protein R1G11.14
8	515.5	23.5	399	2 S41408	lysosomal acid lip
9	513	23.4	399	2 G01416	lysosomal acid lip
10	500.5	22.8	426	2 T20480	hypothetical prote
11	474	21.6	411	2 T22290	hypothetical prote
12	466.5	21.3	411	2 G89074	protein K04A6.5 [i
13	416.5	19.0	559	2 J70949	egg-specific prote
14	392.5	17.9	1585	2 T31611	hypothetical prote
15	356	16.2	443	2 T39540	triglyceride lipas
16	345.5	15.7	467	2 T41053	triglyceride lipas
17	309	14.1	460	2 T39443	probable triglycer
18	293.5	13.4	548	2 S37969	probable triacylg
19	259.5	11.8	413	2 T43170	probable triacygl
20	254	11.6	538	2 S84842	probable membrane
21	237	10.8	573	2 S84754	probable membrane
22	212	9.7	473	2 D86318	protein F15H18.6 [
23	170.5	7.8	509	2 G96766	protein lipase F2P
24	165	7.5	336	2 F83425	probable esterase/
25	147	6.7	460	2 G96764	unknown protein F2
26	121.5	5.5	456	2 E89391	hypothetical prote
27	110.5	5.0	999	2 B70501	hypothetical prote
28	108.5	4.9	1002	2 D87077	probable long-chain
29	107.5	4.9	841	2 A90669	probable enzyme [i

30	107.5	4.9	841	2 D85519	probable enzyme ya
31	107.5	4.9	841	2 C64755	yagx protein - Bsc
32	100	4.6	1216	2 F88473	protein F40H6.5 [i
33	99.5	4.5	798	2 F96714	probable protease
34	99	4.5	323	2 S61927	lipase A precursor
35	98	4.5	465	1 JC1318	triacylglycerol li
36	97.5	4.4	569	2 S75169	urease (EC 3.5.1.5
37	95.5	4.4	461	2 S21223	triacylglycerol li
38	95.5	4.4	467	1 L1DG	triacylglycerol li
39	95	4.3	673	2 T24347	hypothetical prote
40	95	4.3	1648	2 S61654	probable membrane
41	95	4.3	2100	2 T03223	probable polyketid
42	94.5	4.3	835	2 T35763	probable aminopapt
43	94	4.3	519	2 A50726	probable membrane
44	94	4.3	554	2 A47503	epoxide hydrolase
45	94	4.3	569	2 A97648	urease alpha chain

ALIGNMENTS

RESULT 1

E84526

probable lysosomal acid lipase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-feb-2001

C:Accession: E84526

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: R84420; MUID:20083487; PMID:10617137

A:Accession: E84526

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-344 <STO>

A:Cross-references: GB:AE002093; NID:q4585908; PIDN:AAD25569.1; GSPDB:GN00139

C:Genetics:

A:Gene: Atg15230

A:Map position: 2

Query Match 48.1%; Score 1055; DB 2; Length 344;
Best Local Similarity 54.5%; Pred No. 4.2e-82;

Matches 198; Conservative 42; Mismatches 71; Indels 52; Gaps 2;

QY	40	GGSGGLCDOLLLPLGYPCTEHNVTGDLGLSLQHPHGKKAADSTGPPVFLQHLFQ	99
DB	26	GSPVNSLCADLTHPANYSCTEHSIQTKGYILALQKVASLGPRL--QSGPPVLLQHLFM	83
QY	100	GGDTWFINSAEQSLGYILLADNGFDVWGNVGRTRSKGHSTSFVHDKLFDWDSWQELAEY	159
DB	84	AGDVWFNSPKESGLFILADHGFDVWGNVGRTRYSGHVTLSDDTKFWDWSQDLAMY	143
QY	160	DLALMLGYVYVTQSKILYVHSGTIMGALATWPEIVKVMSSAALLCIPISYLDHVSAS	219
DB	144	DLAEMIQYLSISNSKIFLVGHSGQTIMSFAALTQPHVAEMVEAAALLCIPISYLDHVTAP	203
QY	220	FVLRVAMHLDQMLVTMGIHQLNFRSDMGVQIVDSLDCGEHVDCNNLLSAITGENCCFNT	279
DB	204	LVERVFMHLDQ-----	215
QY	280	SRIDYLYLEVEHPSPSTKNLHLFQMRKTFKAYDYGLLGNLRRYGLRPPAFDLSSTPE	339
DB	216	--IEYLDYEPHPSSVKNIHLFQMRKTFKAYDYGVFKNLRTYGLSKPPEILSHIPA	273
QY	340	SUPIWMYGGIDALADVTDVORTIRELGSTPELLYIGDYGHDVFMVSKAKDDVYVDLIR	399
DB	274	SLPWWMGVGGTDGLADVDVDEHTLAELFSSPELLYLEDIGHIDFVIGSSAKEDVYKHMIIQ	333
QY	400	FLR 402	
DB	334	FFR 336	

RESULT 2

T33198

hypothetical protein ZK6.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T33198

R:Wu, X.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid ZK6.

A:Reference number: Z21301

A:Accession: T33198

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-403 <WUX>

A:Cross-references: EMBL:AF067942; PIDN:AA017694.1; GSPDB:GN00023; CESP:ZK6.7

A:Experimental source: strain Bristol N2; clone ZK6

C:Genetics:

A:Gene: CESP:ZK6.7

A:Map position: 5

A:Introns: 29/3; 63/1; 219/3; 319/3; 365/3

C:Superfamily: triacylglycerol lipase, lingual

Query Match 25.4%; Score 557.5; DB 2; Length 403;

Best Local Similarity 32.4%; Pred. No. 1.3e-39;

Matches 124; Conservative 65; Mismatches 153; Indels 41; Gaps 11;

Qy 49 QLLPLGYPCTEHNVTGKDFLLSLQHPHGKNAADSTG--PPVFLQHLFGQGDWTFI 106

Db 31 QIIRWGPAMITVATDDGVILEMHRIPFGKTNVTPNGKRPVVFQWHLGCASSDMMV 90

Qy 107 NSAEQSLGYILLADNGFDWIGNVRGTRWSKGHSTFSVHDKLFDWMSWOELAEYDILLAMLG 166

Db 91 NLPOGAGFLADAGFDWLNMRGNTYSMKHDKLPKSHSAFWDSDWDEMATYDLNAMIN 150

Qy 167 YVYVT-OSKILYVGHSGQTIMGALATMPE--IVKMISAAALCPISYLDHVSAA--SFV 221

Db 151 HVLEVTGQDSVYVGHSGQTLTMSHLSKDDGSAFKIKKFFALAPIGSVKHKGFSLFF 210

Qy 222 LRVAAMHLDQMLVTMGTHQIQL---NFRSDMGVQIVDSLCDGHEVD---CNNLLSAITG-EN 274

Db 211 ANYFSLEFDGWFIDFAGEFIPNNWAMKLAAK---DICGGGLKVEADLCNDVLFITAGPES 267

Qy 275 CFNFSRIDYILEYEPHPSSTKNLHHLFQMTKGTFAKYDYCLGNLRYYGHLRPPAPDL 334

Db 268 DOWNTRVPVYATHDPAGTSTQNIYHWMQVHHGVPAYDAGTKTNKKYGOANPEYDF 327

Qy 335 SSIPESLPIMWYGGGLDALADVTDVQRTIRELGSTPELL-----YIGDYGHI 381

Db 328 TAI-KGTIYLYWSADWLADTPDV-----PDYLLRLNPAIVAQNHLDPYNNHL 376

Qy 382 DFVMSVKAKDDVYVDLIRFLREN 404

Db 377 DPTWGLRAPDDIYRPAIKLCTDD 399

RESULT 3

JC4017

triacylglycerol lipase (BC 3.1.1.3) PGE precursor - bovine

N:Alternate names: pregastric esterase

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: JC4017; S64678

R:Timmermans, M.Y.J.; Teuchy, H.; Kupers, L.P.M.

Gene 147, 239-262, 1994

A:Title: The cDNA sequence encoding bovine pregastric esterase.

A:Reference number: JC4017; MUID:95011625; PMID:7926811

A:Accession: JC4017

A:Molecule type: mRNA

A:Residues: 1-397 <TIM>

A:Cross-references: GB:L26319; NID:g600756; PIDN:AAA57037.1; PID:g600757

A:Experimental source: tongue

Query Match 25.0%; Score 549; DB 2; Length 405;

Best Local Similarity 31.9%; Pred. No. 7.1e-39;

R:Timmermans, M.Y.J.; Reekmans, G.; Teuchy, H.J.H.; Kupers, L.P.M.

Biochem. J. 314, 931-936, 1996

A:Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional

A:Reference number: S64678; MUID:96177869; PMID:8615791

A:Accession: S64678

A:Molecule type: protein

A:Residues: 24-40; 248-253 <TIW>

C:Comment: Pregastric esterase is a major fat-digesting enzyme.

C:Genetics:

A:Gene: pge

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: blocked amino end; carboxylic ester hydrolase; glycoprotein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-397/Product: pregastric esterase #status predicted <MAT>

F:33,270,326/Binding site: carbohydrtate (Asn) (covalent) #status predicted

F:171/Active site: Ser #status predicted

Query Match 25.2%; Score 553; DB 1; Length 397;

Best Local Similarity 34.9%; Pred. No. 3.2e-39;

Matches 130; Conservative 67; Mismatches 146; Indels 30; Gaps 13;

Qy 49 QLLPLGYPCTEHNVTGKDFLLSLQHPHGKNAAD-STGPPVFLQHLFGQGDWTFIN 107

Db 36 QMISYWGYPSEMHKVITADGYILOVYRIPHGKNNANHLGQRPVYVFLQHLGSGSATNWSN 95

Qy 108 SAEQSLGYILLADNGFDWIGNVRGTRWSKGHSTFSVHDKLFDWMSWOELAEYDILLAMLG 167

Db 96 LPKNSLGLFLLADAGYDVLWNGSRGNTWAQEHLYYSPDSPFPAFSDMAEAYDLPSTIDF 155

Qy 168 YVYVT-OSKILYVGHSGQTIMGALATMPEIVYKMISSAALLCPISYLDHVSASFVLRV 225

Db 156 ILRRTGQKLLHYVGHSGQTTIGTAFSTPTLAERIKVYALAPVATVYKTKSLFNKLL 215

Qy 226 AMHL-----DOMLVMTGMIHOLFNRSDMGVQIVDSLCDGHEVD---CNNLLSAITG-ENC 275

Db 216 IPHFLFKIIFGDKFYF---HTF-LEQFLGVE---MCSRETLDVLCNNAFAITGVDNK 267

Qy 276 CFNFSRIDYILEYEPHPSSTKNLHHLFQMTKGTFAKYDYGL-LGNLRRYGHLPAPFDL 334

Db 268 NFNMSRLDVIYAHNPAGTSVQNTLHWQAVKSGKGFQAFDMGAPYQNLHMYHQPPTIYNL 327

Qy 335 SSIPESLPIMWYGGGLDALADVTDVQRTIRELGSTPELLY---IGDYCHIDFVMSVAKD 391

Db 328 TAMNVPVAVSADN-DLLADPDQVDFLLSKLSN---LIYHKEIPNPNYHLDIFWAMDAPQ 382

Qy 392 DVYVDLIRFLREN 404

Db 383 EYNEIVSLMAED 395

RESULT 4

T22675

hypothetical protein F54F3.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T22675

R:Percy, C.; Lloyd, C.

submitted to the EMBL Data Library, September 1996

A:Reference number: Z19598

A:Accession: T22675

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-405 <WIL>

A:Cross-references: EMBL:Z79696; PIDN:CAB01973.1; GSPDB:GN00023; CESP:F54F3.3

A:Experimental source: clone F54F3

C:Genetics:

A:Gene: CESP:F54F3.3

A:Map position: 5

A:Introns: 31/3; 65/1; 274/3; 366/3

C:Superfamily: triacylglycerol lipase, lingual

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C/Accession: G89074
R/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: G89074
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-411 <STO>
A/Cross-references: GB:chr_V; PIDN:AAC48051.1; PID:g1658354; GSPDB:GN00023; CESP:K04A8.5
C/Genetics:
A/Map position: 5
C/Superfamily: triacylglycerol lipase, lingual

	Query Match	21.3%;	Score 466.5;	DB 2;	Length 411;
	Best Local Similarity	32.4%;	Pred. No. 7.8e-32;		
	Matches 118;	Conservative 62;	Mismatches 163;	Indels 21;	Gaps 11;
QY	49	QLLLPLGYPCTEHNVTKDGFLLSLQHHPHGK-----NKAADSTGPPVFLQHGFLFGGGDTW	104		
Db	38	ELIKSWGYSVEIYNTTTKDGFILELHRIPYGREVPTSSDYNNSRPVIFLQHGFLCSSFDW	97		
QY	105	FINSAESGLGVLADNGFDVWIGNVGRTRWSKGHSTFSV-HDKLFDWDSWOELAEYDLLA	163		
Db	98	VANSFPHOSAGFVFADAGFDVWLGNGFRNTVSRKHVSINPKDKPFWDWSWDQISEYDLPA	157		
QY	164	MLGYVYTVT-QSKILYVYGHSGQTINGLAAL-TMPEIVKMISSAALLCPISYLDHVSASFV	221		
Db	158	MIGKALEISGESLYYTGFSLGTMTFAKLTDPKFSRKIKKYFALAPIGSIKHAHGVEL	217		
QY	222	L--RAVAHMLDQMLVTWGIHQLNFRSDMGVQIVDSLQ---DGEHVDCNNLLSAITG---E	273		
Db	218	FLGRHFCKDYEYVKKHGSDLEFGSSLLFKIVKYTCGLFTLLEFCSDITLLFPGTANE	277		
QY	274	NCCFNTSRIDYLYEVEPHSPSTKMLHLHLMFQIMRKGTETAKYDGLLGNLRRYGHRLPPAFD	333		
Db	278	N--NQTRIPYVLAHTAGSSSNVMAHLDQMFSGVGGVPTPDMGEKNLKAYGQKLPPQYN	335		
QY	334	LSSIPESLPIMWGGGLDALADVTDVQRTI-RELGS--TPPELLYIGDYGHIDFVMSVKAK	390		
Db	336	FTGIAD-VPIYLFWSDDDLSTKQDLETLFAQLNSQVVGVSFRINYHLHFIWGTNAA	394		
QY	391	DDVY 394			
Db	395	SQVY 398			

RESULT 13
JT0949
egg-specific protein - silkworm
C/Species: Bombyx mori (silkworm)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
C/Accession: JT0949; A28527
R/Sato, Y.; Yamashita, O.
submitted to JIPID, September 1991
A/Reference number: JT0949
A/Accession: JT0949
A/Molecule type: DNA
A/Residues: 1-559 <SAT>
A/Experimental source: larva
A/Note: This protein is a homotrimer
R/Indrasith, L.S.; Sasaki, T.; Yamashita, O.
J. Biol. Chem. 263, 1045-1051, 1988
A/Title: A unique protease responsible for selective degradation of a yolk protein in Bombyx mori
A/Reference number: A28527; MUID:88087166; PMID:3275655
A/Accession: A28527
A/Molecule type: protein
A/Residues: 19-40,133-144, '146-152;229-248 <IND>
C/Genetics:

A:Map position: 19		Query Match		19.0%; Score 416.5; DB 2; Length 559;	
C:Keywords: egg yolk; homotrimer		Best Local Similarity		31.9%; Pred. No. 2.2e-27;	
F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experi		Matches 12;		Conservative 64; Mismatches 152; Indels 45; Gaps 17;	
F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experi					
Qy	49	QLLLPLGYPCTEHNVETKDGFLLSLQIHIPHGKNAADSTGPPVFLQHLGFGGDTWFINS	108		
Db	194	ELLKKHOYPVEEHTVATDDGVHLTVLRIPPTHQTRDDKKPVALLMHGLLGSADDWLLMG	253		
Qy	109	AEQSLGYILANGDEVDWIGNVGRGTRWSKGHSTFSVHDKL--FWDWSQELAEYDILLAMLG	166		
Db	254	PSKSLAYMLSDAGYDVMGLGNRGNYKRSKH--VSKHPALNDFWKFSNDEIALHDLPAIID	311		
Qy	167	VYIVTVT-QSKTLYVGHSGGTIMGLAALT-MPEIVKMTISSAALLCPISYLDHVSASFVLR	224		
Db	312	HVLDISGOERLUHYIGHSGATFTFALMSQEQSYNEKIVSMHALSPIVYMTNV-RSPLFRM	370		
Qy	225	VA-----MHLDMQ-----LVTMGTHQLN-----FRSDMGVQIVDSLCDGEHVDCCNLL	267		
Db	371	IAPTSTKFYQYIH-DOVGHGAFEPCKHLIETFGGAACREKLCR-----HV-CNNLN	419		
Qy	268	SAITGENCECFNTR--IDYYLEYEPHPSSTKNLHLFOMIRKGFYAKDYDGLLGNLRYG	325		
Db	420	YVISGIN-VYRQDADIVPVMAHLFAGTTSARVMKQYQGNVASHDFRKYNYGAETNMKVYG	478		
Qy	326	HLRPAFDLSIPESLPIMWGYGGDLA-LADVTDVORTIRELGSTPELXYIGDYGH---I	381		
Db	479	ASEPSPYDLSKV--SAPVNL-YHSHDAWLAKPKDVEKLEQLNLPNVKQSFPEPQQHFDTDL	535		
Qy	382	DFVMSVKAKDDVYVDLIRFLREN	404		
Db	536	DFQFSKAPDTVYQKLMEQNMN	558		
RESULT 14					
hypoetical protein Y5088A.g - Caenorhabditis elegans					
C:Species: Caenorhabditis elegans					
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999					
C:Accession: T31611					
R:Steward, C.					
submitted to the EMBL Data Library, September 1999					
A:Reference number: Z21047					
A:Accession: T31611					
A:Status: preliminary; translated from GB/EMBL/DBJ					
A:Molecule type: DNA					
A:Residues: 1-1585 <WLT>					
A:Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB5050.1; CESP:Y5088A.g					
A:Experimental source: clone Y5088A					
C:Genetics:					
A:Gene: CESP:Y5088A.g					
A:Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2					
Query Match		17.9%; Score 392.5; DB 2; Length 1585;			
Best Local Similarity		27.5%; Pred. No. 9.6e-25;			
Matches 114;		Conservative 68; Mismatches 188; Indels 45; Gaps 13			
Qy	11	ALPILLVYVSCGRIVSGASPAAMALRRVSGSGGLCDQLLLPLGYPCTEHNVETKDGFL	70		
Db	4	AVNFFIVFAFKETEETFLDAADTISH-----YGTVEKHVYTTDDGYT	48		
Qy	71	LSLQHIPHGKNAKAA---DSTGPPVFLQHLFGGDTWFINSAEQSLGYILADNGEDVWIG	127		
Db	49	YQLRIPVGRDDRSILGCSKRPVPVFMHGLFGSSYHFLNLNLPSSAAAYIFADAGEDVWL	108		
Qy	128	NVGRTRWSKGHSTFSVHDKLFWDSWQELAEYDILLAMLGYVTVTQ-SKILYVGHSGQTI	186		
Db	109	NIRGTEYGLNHTFSFTNGVYFNWFSLEYSHDYDLRQOIEYFLVLDTRHESLIFYVGHSGQTA	168		

```
QY 187 MCLALTMPEIV-----KMSSAALLCPISYLDHVSASFVLRVAVMHLDQMLVTWG 237
Db 169 VMFARLAEDVTWQSKIRVFFALPTAGFMKLPFTLLLENYLQALIQFALDG---RFG 225
QY 238 IHLNFRSDMGVQIVDSLCDGSHVD--CNLLSAITGENCC--FNTSRIDYVLEYEPHPS 293
Db 226 ILPVEIPRAISKFAD-FCSSKFFFLCSAGFKVAAGIETLQOVNDSRIPIILSHFPSAT 284
QY 294 STKNLHLFQMIKRTGTFKAYDYGLGNLRRYRGLRPPAFDLSSIPESLPIMWGYGGLDAL 353
Db 285 STLLNLLHWQIPKHYHELRLDLGTARNLIAYGQKAPRLEIGNIIAQITLY--FSKDDQI 342
QY 354 ADVTDVQRTI-RELGSTPELLYIGD---YGHIDFVMSVKAKDDVYVDLI-RELRE 403
Db 343 TDEVDVREIIMKMG--FGLIESYDLHDHFTFDFILGLRATDEVYKPIVIYRIYKE 395

RESULT 15
T39540
triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39540
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
submitted to the EMBL Data Library, February 1998
A:Reference number: 221862
A:Accession: T39540
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-443 <WOO>
A:Cross-references: EMBL:AL021748; PIDN:CAA16863.1; GSPDB:GN00067; SPDB:SPBC16A3.12c
A:Experimental source: strain 972h-; cosmid c16A3
C:Genetics:
A:Gene: SPDB:SPBC16A3.12c
A:Map position: 2
A:Introns: 11/2: 152/3: 282/1: 398/3
C:Superfamily: triacylglycerol lipase, lingual

Query Match 16.2%; Score 356; DB 2; Length 443;
Best Local Similarity 28.3%; Pred. NO. 2.3e-22;
Matches 108; Conservative 75; Mismatches 149; Indels 50; Gaps 16;

QY 55 GYPCTEHNVEYKDGFLLSLQHIPHGKNAADSTGPPVFLQGLFOGGDTWF-INSABOSL 113
Db 83 GYRVEEHLVRTQDNFILCLHRIHPKQ--SQHKREVVYCHHGLMTNSELWAVNESERSL 140
QY 114 GYLADNGFDVWIGNVGRTRWSKGHSTFSVHDKLFWDKSWQELAEYDILLAMLGYYVTVT- 172
Db 141 PVFLIESGTDVWGLNNGRNKYSRKHIITYKPRDEEFNFSLDDMAMFDIPDTVDYILRETG 200
QY 173 QSKILYVGHSGQTIMGLAALTM-PEIVKMISSAALLCP-----ISYLDHVSASF 220
Db 201 REKLVNIGFSQGTQAAMAALSINPDLNDKVNIFIGLAPAYAPKGFSNFYVDYIVKVNPKI 260
QY 221 VLRVAMHLDQMLVTMGTHQLNFRSDMGVQIVD-SLCDGEHVDCNNLLSAITGENCCPNT 279
Db 261 MYHLEGRRC--LLPSVTFWQNICYPPIEVKIVDVSLSL-----KILFNWDLNLSLNQ 309
QY 280 SRIDYVLEYEPHPSSTKNLHLFQMIKRTGTFKAYDYGLGNLRRYG--HLRPPAFDLSSI 337
Db 310 KLCGFAHLYS--FSSVKSVVHWLQIKNKTFQLYDDD-MALLAGYGSRRHYQVPLEPTNNI 366
QY 338 PESLPIMWGYGGLDALADVTQVORTIRELGSTPELLYIGDYGHIDFVMSVKAKDDVY--- 394
Db 367 --KCPMLILWGGKDTLINN-EVMRT--ALPPHAKVSIATYEHLDLFWGQDVKEEVPV 421
QY 395 VDLIR-----FLRENGWH 407
Db 422 IDALKHSHSLGRKAKHEVKQNGFH 443
```

Search completed: December 30, 2002, 20:47:19
Job time : 53 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	553	25.2	397	1	LIPG_BOVIN	Q29458 bos taurus
2	546	24.9	398	1	LIPG_HUMAN	P07098 homo sapien
3	539.5	24.6	395	1	LIPG_RAT	P04634 rattus norv
4	529	24.1	398	1	LIPG_CANFA	P80035 canis famill
5	515.5	23.5	399	1	LICH_HUMAN	P38571 homo sapien
6	513.5	23.4	394	1	LIP3_DROME	A6108 drosophila
7	503	22.9	397	1	LICH_MOUSE	Q9Z0m5 mus muscullu
8	486	22.2	397	1	LICH_RAT	Q64194 rattus norv
9	427	19.5	439	1	LIP1_DROME	Q64107 drosophila
10	293.5	13.4	548	1	TG11_YEAST	P34163 saccharomyce
11	107.5	4.9	841	1	YAGX_ECOLI	P77802 escherichia
12	100	4.6	1216	1	YPX5_CAEEL	Q90277 caenorhabdi
13	97.5	4.4	569	1	UR1L_SYNY3	P23061 synechocyst
14	95.5	4.4	461	1	LIP1_HORSE	P29183 equus cabal
15	94	4.3	465	1	LIPP_RABIT	Q02157 ryctolagus
16	94	4.3	554	1	HVES_RAT	P80299 rattus norv
17	93.5	4.3	467	1	LIP1_CANFA	P06857 canis famill
18	93.5	4.3	1167	1	CIJ3_BACTU	Q45738 bacillus th
19	93.5	4.3	1680	1	CS3_MOUSE	P06684 mus muscullu
20	92	4.2	239	1	SFSA_AGRT5	P58429 agrobacteri
21	91	4.1	465	1	LIPP_CAVPO	P50903 cavia porce
22	91	4.1	569	1	URE1_BACSU	P77837 bacillus su
23	90.5	4.1	465	1	LIPP_HUMAN	P16233 homo sapien
24	90	4.1	1676	1	CO3_HUMAN	P01031 homo sapien
25	89.5	4.1	445	1	SCRC_RABIT	A46502 oryctolagus
26	88.5	4.0	518	1	YOAE_ECOLI	P76262 escherichia
27	88.5	4.0	669	1	GALC_CANFA	P54804 canis famill
28	88.5	4.0	1876	1	P3SA_MYCTU	Q10977 mycobacteri
29	88	4.0	457	1	LIPP_MYOCO	Q64425 myocastor c
30	88	4.0	972	1	KFMS_HUMAN	P07333 homo sapien
31	86.5	3.9	281	1	YN07_MYCTU	Q50658 mycobacteri
32	86.5	3.9	307	1	PYRB_STRPN	Q97qe2 streptococc
33	86	3.9	519	1	DPY5_MOUSE	Q9eqf5 mus muscullu

Db 36 QMISYWGYPSEMHKVIITADGYILOVYRIPHGKNNANHLGQRPVVFLOHGLGSAINWISN 95
Qy 108 SAEQSLGYILLADNGFDWIGNVGRTRWSKGHSTESVHDKLFDWMSWOELAEYDILLAMLY 167
Db 96 LPKNSLGLFLADAGYDVLGNSRGTWAQEHLYSPSPFEWAFSDEMAEYDLPSTIDF 155
Qy 168 VYTVT-QSKILYVGHSGQTMGLAAL-TMPEIVKMISAAALLCPISYLDHVSASFVLRV 225
Db 156 ILRRTGQKLLHYVGHSGQTMGLAAL-TMPEIVKMISAAALLCPISYLDHVSASFVLRV 215
Qy 226 AMHL-----DQMLVTMGHQLNFRSDMGVQIVDLSLDCGEHVD--CNNLLSAITG-ENC 275
Db 216 IPHFLKIFIDKMFY---HTF-LEQFLGVE----MCSRETLVLCKNALFAITGVNDK 267
Qy 276 CFNFSRIDYLYEPHPSSTKNLHLMQIRKGFYAKYDYL-LGNLRRYGHLPAPFDL 334
Db 268 NFNSRLDVIYAHNPAGTSVONTLHWQAVKSGKFAQFDWGPQNLHMYHQPTPPYINL 327
Qy 335 SSIPESLPIMWGYGGLDALADVTQVIRELGSTPELly---IGDYCHIDFVMSVRAKD 391
Db 328 TAMNVPVIAWSADN-DLLADPQDVDFLLSKLSN--LIYHKEIPNYNHLDFIAMDAPQ 382
Qy 392 DYYVDLIRFLREN 404
Db 383-EVYNEIVSLMAED 395

RESULT 2
LIPG_HUMAN
ID LIPG_HUMAN STANDARD; PRT; 398 AA.
AC P07098;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
lipase) (GL).
GN LIPG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299724; PubMed=3304425;
RA Bodmer M.W., Angel S., Yarranton G.T., Harris T.J.R., Lyons A.,
RA King D.J., Pieroni G., Riviere C., Verger R., Lowe P.A.;
RT "Molecular cloning of a human gastric lipase and expression of the
RT enzyme in yeast."
RL Biochim. Biophys. Acta 909:237-244(1987).
RN [2]
RP SEQUENCE OF 20-45.
RX MEDLINE=89325292; PubMed=2753032;
RA Bernbaeck S., Blaeckberg L.;
RT "Human gastric lipase. The N-terminal tetrapeptide is essential for
RT lipid binding and lipase activity."
RL Eur. J. Biochem. 182:495-499(1989).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=99287897; PubMed=10358049;
RA Roussel A., Canaan S., Egloff M.P., Riviere M., Dupuis L., Verger R.,
RA Cambillau C.;
RT "Crystal structure of human gastric lipase and model of lysosomal
RT acid lipase, two lipolytic enzymes of medical interest."
RL J. Biol. Chem. 274:16995-17002(1999).
CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
Db EMBL; X05997; CAA29413.1; -
Db EMBL; X05997; CAA29414.1; ALT_INIT.
Db EMBL; A01046; CAA00125.1; -
Db EMBL; A12714; CAA01053.1; -
Db PIR; S04942; S04942.
Db PIR; S07145; S07145.
Db PDB; 1HLG; 15-MAR-00.
Db Genew; HGNC:6622; LIPF.
Db MIM; 601980; -
Db InterPro; IPR000073; Abhydrolase.
Db InterPro; IPR000734; Lipase.
Db InterPro; IPR000379; Ser_estrs_site.
Db Pfam; PF00561; abhydrolase_1.
Db PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Polymorphism;
KW 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
FT ACT_SITE 172 172 CHARGE RELAY SYSTEM.
FT ACT_SITE 343 343 CHARGE RELAY SYSTEM.
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM.
FT DISULFID 246 255
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .).
FT VARIANT 161 161 T->A (IN DBSNP:814628).
FT /FTID=VAR_011947.
SQ SEQUENCE 398 AA; 45237 MW; CD3EE1621C014F0F CRC64;
Query Match 24.98; Score 546; DB 1; Length 398;
Best Local Similarity 33.98; Pred. No. 1.2e-39;
Matches 130; Conservative 70; Mismatches 132; Indels 52; Gaps 13;
Qy 49 QLLPLPGYPCPEHNVETKDGFLLSLQHIPHGKNAADSTG--PPVFLQHLGFGQGDWTFI 106
Db 37 QMITYWGYPNEEYEVWTEGYILEVNRIPYCK-KNSGNTGQRPVVFLOHGLLASATNWS 95
Qy 107 NSAEQSLGYILLADNGFDWIGNVGRTRWSKGHSTESVHDKLFDWMSWOELAEYDILLAMLY 166
Db 96 NLPNNSLAFLADAGYDVLGNSRGTWAQEHLYSPSPFEWAFSDEMAEYDLPATID 155
Qy 167 VYTVT-QSKILYVGHSGQTMGLAAL-TMPEIVKMISAAALLCPISYLDHVSASF-VLR 223
Db 156 FIVKTKGQKLLHYVGHSGQTMGLAAL-TMPEIVKMISAAALLCPISYLDHVSASF-VLR 215
Qy 224 AVAMHL-----DQMLVTMGHQLNFRSDMGVQIVDLSLDCGEHVD--CNN 265
Db 216 FVPSQLFKFIDGKIFYPHNFDDQLAT-----EYCSREMLNLLCSN 257
Qy 266 LLSAITG-ENCCFNTRIDYLYEPHPSSTKNLHLMQIRKGFYAKYDYL-LGNLRR 323
Db 258 ALFIICGFDKSNFNTSRDLVLSHNPAGTSVQNMFHMTQAVKSGKFAQYDWGSPVQNRMH 317
Qy 324 YGHLRPPAFDLSLPESLPIMWGYGGLDALADVTQVIRELGSTPELly---IGDYCH 380
Db 318 YDQSQPPYINVTAMNVPVIAW--NGGKDLLADPQDVDFLLSKLSN--LIYHKEIPNYNHL 372
Qy 381 IDFVMSVRAKDDYVVDLIRFLREN 404
Db 373 LDFIAMDAPQEVYNDIVSMISED 396
RESULT 3
LIPG_RAT
ID LIPG_RAT STANDARD; PRT; 395 AA.
AC P04634;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)


```
FT VARIANT 16 16 P -> T.
FT VARIANT 129 129 H -> P (IN CESD).
FT VARIANT 129 129 H -> P (IN CESD).
FT VARIANT 129 129 H -> R (IN CESD).
FT VARIANT 200 200 L -> P (IN CESD AND WD).
FT VARIANT 23 23 G -> R (IN REF. 3).
FT VARIANT 29 29 V -> L (IN REF. 4).
SQ SEQUENCE 399 AA: 45415 MW: 55F27391306B609A CRC64;

Query Match 23.5%; Score 515.5; DB 1; Length 399;
Best Local Similarity 32.3%; Pred. No. 5.4e-37;
Matches 131; Conservative 59; Mismatches 143; Indels 73; Gaps 12;

QY 42 GSGGL-----CDQLLLPLGVPCTEHNVEKDGFLLSLQHPHG-KNKAADSTGP 89
DB 21 GSGGLTAVDPETNNVSEIISYWGFPSEYIYVEDYGVILCLNRLPHGRKKNHSDGPKP 80
QY 90 PVFLQHLGFGQGDWTFINSAEQSLGYILADNGFDVWIGNVGRTRWSKGHSTFVHDKLFW 149
DB 81 VVFLQHLGLLADSSNVNTLANSLSGIFLADAGFDVWGMGRGNTWSRKHKTLSVSDQEFW 140
QY 150 DMSWELAEYDLMLGYYVTVT-OSKILYVGHSGCTINGLAALT-MPIVVKMISAAALL 207
DB 141 AFSYDEMAKYDLPASINFILNKTGQEQVYVGHSGCTTGTFATFQIPELAKRIRKMFAL 200
QY 208 CPISYLDHVSASFVLRVAVM-----HLDQMLVTWGIHOLFNRSDMGVQIVDSLCD- 257
DB 201 GPVA-----SVAFCSPMAKGLRPLDHL-----IKDLFGDKFEL 234
QY 258 -----GEHVDCCNNLSAITGNCFC-----NTSRIDYILEYEPHPSSTKNL 298
DB 235 POSAFLKWLGTWCHVILKELCG-NLCELLCGFNERNLNSRVDVYTHSPAGTSVQNM 293
QY 299 HHLFQMRKGTFAKYDYLGLG-NLRRYGHLPAPFDLSSIPESLPIMWGMGGLDALADVT 357
DB 294 LHWQAVKFKQFAQDWGSSAKNYFHNQSYPTNVKMDLVPATVWS--GGHDWLADVY 351
QY 358 DVQRTIRELGSPPELLYIGDYGHIDFVMSKAKDDVYVDLIRFLE 403
DB 352 DVNILLTQITNLVFHESIPWEHLOFVGLDAPWRLNKKIINLMRK 397

RESULT 6
ID LIP3 DROME STANDARD; PRT; 394 AA.
AC O46108;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Lipase 3 precursor (EC 3.1.1.-).
GN LIP3 OR C68823.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=9827315; PubMed=9566193;
RA Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RT tissue and developmental specific expression.";
RL J. Mol. Biol. 276:877-885(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
```

```
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis J.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flossler C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- TISSUE SPECIFICITY: FAT BODY.
CC -!- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y14367; CAA74737.1;
CC EMBL: AE003699; AAF54935.1;
CC FlyBase: FBgn0023495; Lip3.
CC InterPro: IPR000073; Abhydrolase.
CC InterPro: IPR000734; Lipase.
CC InterPro: IPR000379; Ser_estr_site.
CC Pfam: PF00561; abhydrolase_1.
CC PROSITE: PS0120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Signal; Glycoprotein.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 394 LIPASE 3.
FT ACT_SITE 164 164 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 369 369 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 131 131 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 394 AA: 44901 MW: A718D1D743673802 CRC64;

Query Match 23.4%; Score 513.5; DB 1; Length 394;
Best Local Similarity 31.4%; Pred. No. 7.9e-37;
Matches 128; Conservative 78; Mismatches 162; Indels 39; Gaps 14;

QY 10 GALPILLLVWSCGGRIVSGASPAALRRVSGSGGLCDQLLLPLGYPCTEHNVEKDFG 69
DB 4 GALKVTILLVG-LGLVLGSRPISD-----CCGRIEDDGYPMERHEVVTSDNY 50
```


RX MEDLINE=96129534; PubMed=8576647;
RA Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,
RT Yoshida H., Osame M.;
RT "Cloning of rat lysosomal acid lipase cDNA and identification of the
RT mutation in the rat model of Wolman's disease.";
RL J. Lipid Res. 36:2212-2218(1995).
CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
CC acid.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S81497; AAB36043.2; -;
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; FALSE_NEG.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
FT SIGNAL 1 19
FT CHAIN 20 397
FT ACT_SITE 172 172
FT ACT_SITE 372 372
FT CARBOHYD 34 34
FT CARBOHYD 99 99
FT CARBOHYD 159 159
FT CARBOHYD 271 271
FT CARBOHYD 319 319
SQ SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;
Query Match 22.2%; Score 486; DB 1; Length 397;
Best Local Similarity 33.4%; Pred. No. 1.9e-34;
Matches 124; Conservative 65; Mismatches 154; Indels 28; Gaps 14;
QY 49 QLLPLPGYCTEHNVTGDKGFLLSLQHPHG-KNKAADSTGPPVFLQ--HGLFQGGDTWF 105
DB 37 EIIMHWGYP--EHSVOTGDCYILGVHVRPHGRKNQFDKGPVYVQLWRHGLADSSNV 94
QY 106 INSAEOSLYIADNAGFDYWIGNVTRTSRKGHSTSVHDKLFDWDSWELABYDLLAML 165
DB 95 TNIDNNSLGFIILADAGFDYVMGNSRGNTSRKHKTLSVSQDEYWFASFDEMAKYDLPASI 154
QY 166 GYVYVTQSKILY-VGHSOGTIMGALALT-MPEIVKMISSAALLCPISYLDHVSASFVLR 223
DB 155 NVILNKTGQELNVGHSOGTIGFTAFSOMPELAKVMFFALAPVLSNLFASGPMVKL 214
QY 224 AVAMHLDQMLV--TMGIHQLNFRSDMGVQIVDSLCDGEHVD-----CNLLLSAITGEN-CC 276
DB 215 G---RLPDLLEDLFQKQFLQPSAMVKKLSTHIC--THVMKELCANIFFLICGFNEKN 269
QY 277 FNTSRDYLYEYEPHPSSTKNLHLEFQMTKGTFAKYDYGLLG-NLRRYGHURLAPPADLS 335
DB 270 LNNRSVDVYTHCPAGTSVQNVHVTQVYKYLQAFDMGSSDKNVFHYNQSYPLYSIK 329
QY 336 SIPESLPIWVGGLDALADVTDVORTIRELGSTPELLY---IGDYGHIDFVMSVAKDD 392
DB 330 DM--QLPTALWGGKDWLADTSDINILLTEI---PTLVYHKNIPEMDHLDFTWGLDAPWR 384

QY 393 VYVDLIRFLRE 403
DB 385 LYNEVVSLMKK 395
RESULT 9
LIPL_DROME
ID LIPL_DROME STANDARD; PRT; 439 AA.
AC O46107: Q9VKR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipase 1 precursor (EC 3.1.1.-).
GN LIPL OR CG7279.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE OF 7-433 FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98227315; PubMed=9566193;
RA Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RT tissue and developmental specific expression.";
RL J. Mol. Biol. 276:877-885(1998).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,
RA Sutton G.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Abayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,
RA Burkitt K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swigras R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -!- TISSUE SPECIFICITY: OVARIES.
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO

```
CC -----
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y14366; CAA74736.1; -.
DR EMBL; AE003629; AAF52994.1; -.
DR FlyBase; FBgn0033496; Lip1.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; FALSE_NEG.
KW Hydrolase; Lipid degradation; Signal; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 439
FT DOMAIN 30 44
FT ACT_SITE 197 197
FT ACT_SITE 393 393
FT CARBOHYD 124 124
FT CARBOHYD 151 151
FT CARBOHYD 346 346
FT CARBOHYD 379 379
FT CARBOHYD 426 426
FT CONFLICT 10 10
FT CONFLICT 213 213
FT CONFLICT 412 412
FT SEQUENCE 439 AA; 50660 MW; 9E32E20BEAE93E3F CRC64;

Query Match 19.58; Score 427; DB 1; Length 439;
Best Local Similarity 30.18; Pred. No. 2.7e-157;
Matches 110; Conservative 67; Mismatches 157; Indels 32; Gaps 12;

Qy 48 DQLLLPLGYPCTEHNVERKDGFLSLQHPHGKKAADSTGPPVFLQHLFGQGGDTWFIN 107
Db 69 DKLIAGYGESEVHHVTTEDGYLTMHRI---RKQGA---PPFLQHLGVDSSAGFVVM 121
Qy 108 SAEOSLGVLADNGFDVWGNVGRTRWSKGHSTSVHDKLFDWDSWQELAEYDILLMLGY 167
Db 122 GPNVSLAYLLADHNDVWLGAGNRNRYSRNHTLLDPDESKFDFSWHEIGMYDLPAMDH 181
Qy 168 YTVTQ-SKILYVGHSGQTIMGALALT-PEIVKMISAAALLCPISYLDHVSASFVLRAV 225
Db 182 VLKVTGFPKLYHAGHSQGTCTFFVCMRPAYNKDVSMQALAPAYAKETEDHPYIRAI 241
Qy 226 AMHLDQMLVTGHIOL---NFRSDMGVQIVDSLC---DGEHVDCNNLSAITGENCC-F 277
Db 242 SLYPNS-LVGSSIREWFNGEPR-----FLCRMTETEL-CIEAVFGIVGRNWF 290
Qy 278 NPSRIDYLEYEPHPSSTKNLHLLFQMIKGTFAKYDGLLGNLRLYGHLPAPDLSSI 337
Db 291 NRKMFPVILGHYPAGAAQVKHFTIILKSGRFAPYSYSSNKNMQLYRDLPLPPRYNLSIV 350
Qy 338 PSLPIWMGYGGGLDALADVTQVORTIRELST--PELLYIGDYGHIDFVMSKADDDVYV 395
Db 351 --TVPTFVYSNDLLCHPKDQVSMCDDLGNVTGKLYLPQKEFNHMDLWADIVRKMLYR 408
Qy 396 DLIRFL 401
Db 409 RMLQVL 414

RESULT 10
ID TGLL YEAST
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
```

```
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triglyceride lipase-cholesterol esterase (EC 3.1.1.-).
GN TGL1 OR YKL140W OR YKL5.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92245761; PubMed=1574929;
RA Abraham P.R., Mulder A., Van't Riet J., Planta R.J., Raue H.A.;
RT "Molecular cloning and physical analysis of an 8.2 kb segment of
RT chromosome XI of Saccharomyces cerevisiae reveals five tightly linked
RT genes.";
RL Yeast 8:227-238(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Red M.R., Xu G., Kirchrath L., Fritz C., Keuchel H., Hollenberg C.P.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE AB HYDROLASE SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z25464; CAA80958.1; -.
DR EMBL; Z28140; CAA81981.1; -.
DR PIR; S37969; S37969.
DR PIR; S39000; S39000.
DR SGD; S0001623; TGL1.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
KW Hydrolase; Lipid degradation.
SQ SEQUENCE 548 AA; 62979 MW; 32D1F230701CB083 CRC64;

Query Match 13.48; Score 293.5; DB 1; Length 548;
Best Local Similarity 24.88; Pred. No. 1.2e-17;
Matches 95; Conservative 67; Mismatches 136; Indels 85; Gaps 16;

Qy 60 EHNVTETKDGFLSLQHPH-GKNKAADSTGPPVFLQHLFGQGGDTWFIN-SAEOSLGYL 117
Db 79 DHLVETEDNYILTHRIPIPPISKRNKVV---VYLHGLLSCSDVWCCNIERHKNLPVL 135
Qy 118 ADNGFDVWGNVGRTRWSKGHSTSVHDKLFDWDSWQELAEYDILLMLGYVTVTQ-SKI 176
Db 136 HDLGYDVWGNNGNKYSTAHLNKKPKSKNKEWDFSIDFAFDIPNSIEFIDITDKVKV 195
Qy 177 LYVGHSGQTIMGALALT-PEIVKMISAAALLCPISYLDHVSASFVLRAVAMHLDQMLVTM 236
Db 196 ICIGFSQSAQMFASFUSE-----KLARKVSHFIAIAPAM-----TPK 234
Qy 237 GIHQLNFRSDMGVQIVDSLCDE-----HVDCCNNLSAITG 272
Db 235 GLHN-----RIVDTLAKSPGMYLFFGRKIVLPSPAVIQWTLHTLFLNLCIDIAN 285
Qy 273 ENCCFN-----TSRIDYLEYEPHPSSTKNLHLLFQMIKGTFAKYDY--GLLGNLR 322
Db 286 K-ILFNWKSFNILPQKTIASYAKLY-STTSVKVSIWHFQILRSQKQFOMFEESDNNLSLT 343
Qy 323 RYGHLPAPDLSSIPE----SLPIWMGYGGGLDALADVTQVORTIRELGSYPELLYIGDY 378
Db 344 R-----PYQIANPPTTNIKIPILLIYGGIDSLVDI-DVWKNLPPNSVFD-VKVDNY 394
Qy 379 GHIDFVMSKADDDVYVDLIRFL 401
Db 395 EHLDLGWKADDTLVIAKVLRFI 417
```


RESULT 13
ID UREL_SYN3 STANDARD; PRT; 569 AA.
AC P73061;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN UREAC OR SLL1750.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: (ALPHA, BETA, GAMMA)(3) (BY SIMILARITY).
CC -!- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
ligands for the two nickel ions (By similarity).
CC -!- SIMILARITY: BELONGS TO THE UREASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
EMBL; D90903; BAA17083.1; .
DR HSSP; P18314; 1FWE.
DR MEROPS; M38.0NW; .
DR InterPro; IPR001924; UreaseA.
DR Pfam; PF00449; urease; 1.
DR Pfam; PF02802; urease; 1.
DR ProDom; PD002467; UreaseA; 1.
DR PROSITE; PS00145; UREASE_2; 1.
DR PROSITE; PS01120; UREASE_1; 1.
KW Hydrolase; Metal-binding; Nickel; Complete proteome.
FT METAL 136 136 NICKEL 2 (BY SIMILARITY).
FT METAL 138 138 NICKEL 2 (BY SIMILARITY).
FT METAL 219 219 NICKEL 1 AND 2 (BY SIMILARITY).
FT METAL 248 248 NICKEL 1 (BY SIMILARITY).
FT METAL 274 274 NICKEL 1 (BY SIMILARITY).
FT METAL 362 362 NICKEL 2 (BY SIMILARITY).
FT ACT_SITE 322 322 BY SIMILARITY.
SQ SEQUENCE 569 AA; 61037 MW; 1C720A254335D7FD CRC64;
Query Match 4.4%; Score 97.5; DB 1; Length 569;
Best Local Similarity 19.6%; Pred. No. 1.1; 135; Indels 105; Gaps 16;
Matches 71; Conservative 51; Mismatches 51;
QY 77 PHGKNKADSTGPPVFLQHLGFGDGTWTFINSAEQSLGYLADNGFDVWIGNVGRTRWSK 136
DB 163 PRAGTKATCTT-PGAWNTHRMLOADGFPIN-----LGLFGKNG-----SQPAALAE 209
QY 137 GHSTFSPVHDKLFWDMSQBELA-----EYDL-----LAMLGYVYT---VTQSKI 176
DB 210 QIKAGAIGLKLHEDWCTTPAIDNCLGVAEDYDVQVAIHDTLNEAGFVETTTAALKNRV 269
QY 177 LVVGHSGQTIMGALATMTPEIVKMISSALLCPISVLHDVSAFVLRAVAMHLDQMLVTM 236
DB 270 IHTYHTGAGGGGHA-----PDIKICGELNVL-PSS--TNPTRPYTINTLEEHLDMLVCH 322

QY 237 GIHQLNFRSDMGVQIVDSLCDGHEVDCNNLLSAI----- 270
DB 323 HLHR-NIPED--VAFAESRIRRETIAEDILHDGAFSIISSDSQAMGRVGEVICTWQT 379
QY 271 -----TGNCCFNFSRIDYLEYEPHPSSTKNLHLFQMIKRGTF----- 311
DB 380 AHKMKVQRGQLPGETGNNDNLRAKR--YVAKYITINPAITHGISEEIGSVVEGKLADLCW 437
QY 312 -----KYDYGLLGNLRRYGHLPAPAFDLSSSIPESLPIWN-----GYGGLDALADYTDV 359
DB 438 SPAFFGVKPELVIKGGIVAVQAMCDAN--ASITPPQVHMQPMFANYGGAIAATSVTFF 494
QY 360 QR 361
DB 495 SQ 496
RESULT 14
LIPP_HORSE
ID LIPP_HORSE STANDARD; PRT; 461 AA.
AC P29183;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, pancreatic precursor (EC 3.1.1.3) (Pancreatic
lipase) (PL) (Fragment).
GN PNLIIP.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92267022; PubMed=1587279;
RA Kerfelec B., Foglizzo E., Bonicel J., Bougis P.E., Chapus C.;
RT "Sequence of horse pancreatic lipase as determined by protein and
cDNA sequencing. Implications for p-nitrophenyl acetate hydrolysis by
pancreatic lipases";
RL Eur. J. Biochem. 206:279-287(1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=94238688; PubMed=8182745;
RA Bourne Y., Martinez C., Kerfelec B., Lombardo D., Chapus C.,
Cambillau C.;
RT "Horse pancreatic lipase. The crystal structure refined at 2.3-A
resolution";
RL J. Mol. Biol. 238:709-732(1994).
CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
EMBL; X66218; CAA46961.1; .
DR PIR; S21251; S21251.
DR PIR; S21223; S21223.
DR PDB; 1HPL; 31-MAY-94.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR000379; Serestrifite.
DR Pfam; PF01477; Lipase; 1.
DR Pfam; PF01477; PLAT; 1.
DR SMART; SM00308; LH2; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Pancreas; Signal; 3D-structure.


```

FT NON_TER 1 1
FT SIGNAL <1 12
FT CHAIN 13 461
FT ACT_SITE 165 165
FT ACT_SITE 189 189
FT ACT_SITE 276 276
FT DISULFID 16 22
FT DISULFID 103 114
FT DISULFID 250 274
FT DISULFID 298 309
FT DISULFID 312 317
FT DISULFID 445 461
FT MUTAGEN 165 165
FT SQ SEQUENCE 461 AA: 50921 MW: 382F33F3CE446738 CRC64;

Query Match
Best Local Similarity 4.4%; Score 95.5; DB 1; Length 461;
Matches 94; Conservative 56; Mismatches 155; Indels 149; Gaps 25;

QY 57 PCTEHNVTGDKGFLLSLQHIPIHG-KNKAAD-----STGPPV-FLOHGLFQGGDTWF 105
Db 41 PWSPEKVNR--FLLYTNENPNFQEIADPSTIQSSNFNTGKTRFIHG-----F 90
QY 106 INSAEQLGYILADNGFDVWIGNVRGTRMSKGHSTFSVHDKLFWDMWQELAEYDILL--A 163
Db 91 IDKGESWLSMCQNMFKVESVNCICVDMKSGSRT-----AYSQASQNVRIVGA 139
QY 164 MLGYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 195
Db 140 EYAVLVGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 199
QY 196 EIVKMISSAALLCPISYLDHV-----SASFLRAVAMHLDQMLVTCGIHQLNFRSDMGVQI 251
Db 200 ELVRLDPSDAQFVDVIHTDIAPFIFNLFGMSQTAGHLD-FFPNGGKMPGCKQKVLISOI 258
QY 252 VDSLCDG-----EIVDCNNLLS-----AITGENC-----CFNTS----- 280
Db 259 VD--IDGIWQGRDFAACNHLRSKYKYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 316
QY 281 --RIDYLYEYEPHPSSTKNLHLFQIMIR--KGTFAKYD-----G 316
Db 317 CPOMGHYA--DRFPQRTGKVGOLFVYLTNGDASNFARWRYVDVTLGSKKVTGHVLSVLF 374
QY 317 LLGNLRRY----GHLRPP-----AFD-----LSSIPESLPIWMYGGGLDALADVTDVQRTI 363
Db 375 NKGNSROYEIFOGTLKPDNTYNEEDSDVEVGDLKVKFIW--YNNV-----INLTL 424
QY 364 RELGSTPELLYIGDYGHIDFVMSVRAKDDVYVDL 397
Db 425 PKVGASKITVERNDGVSFNFCEETVREDVLTLL 458

RESULT 15
LIPP_RABBIT STANDARD: PRT: 465 AA.
AC Q02157;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, pancreatic precursor (BC 3.1.1.3) (Pancreatic lipase) (PL).
DE GN PNLIP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93075235; PubMed=1445366;
RA Aleman-Gomez J.A., Colwell N.S., Sasser T., Kumar V.B.;
RT "Molecular cloning and characterization of rabbit pancreatic triglyceride lipase.";

```

```

RL Biochem. Biophys. Res. Commun. 188:964-971(1992).
CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M99365; AAA31489.1; -.
DR HSP; P29183; 1HPL..
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR00379; Ser_estrs_site.
DR Pfam; PF00151; Lipase; 1.
DR Pfam; PF01477; PLAT; 1.
DR PRINTS; PR00821; TAGLIPASE.
DR SMART; SM00308; LH2; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Pancreas; Signal.
FT SIGNAL 1 16
FT CHAIN 17 465
FT ACT_SITE 170 170
FT ACT_SITE 194 194
FT ACT_SITE 281 281
FT DISULFID 20 26
FT DISULFID 108 119
FT DISULFID 255 279
FT DISULFID 303 314
FT DISULFID 321 321
FT DISULFID 449 465
SQ SEQUENCE 465 AA: 51161 MW: 47F53275997BBA19 CRC64;

Query Match
Best Local Similarity 4.3%; Score 94; DB 1; Length 465;
Matches 106; Conservative 57; Mismatches 168; Indels 182; Gaps 27;

QY 11 ALPLLVLVVS-----CCGRIVSGASPAALRRVSGSGGLCDQLLLPLGVPCTE 60
Db 6 ALPLLGVAGLEVCEYERLGCFCGNRI-----PWSGGTLERPFSTL--PSTP 49
QY 61 HNVETKDFGLLSLQHIPIH-----CKNKAADSTGPPVFLQHLGF-OGGDTWFI 106
Db 50 KIVNTR--FLLYTNENPNNFQEIADASTIRGSNFTDR--KTRFTIHGFTDKGEENWLS 105
QY 107 NSAEQSLGYILADNGFDVWIGNVRGTRMSKGHSTFSVHDKLFWDMWSQEL----AEY--- 159
Db 106 N-----LCENLFQVETVNCICVDMKSGSRT-----TYQATQNIIRIVGAEVAYL 149
QY 160 --DLLAMLGYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 196
Db 150 VGTLOSSLYG----SFSNIHVIGHSLGAHAAGEVGRRTNGTIGRTGLDPAEPYFGQTPE 205
QY 197 IVKMISSAALLCPISYLDHV-----SASFVLRVAVAMHLDQMLVTMGTHQLNFRSDMGVQIV 252
Db 206 IVRLDPSDAQFVDVIHTDAAPVNPVLGFGMSQTVGHLD-FFPNGGKMPGCKQKVLISOIV 264
QY 253 D--SLCDG--EIVDCNNLLS-----AITGENC-----CFNTS-----RI 282
Db 265 DTNGWEGTRDFVACNHLRSKYKYADSVINPNNGFAGFSCASYAFSANKKFCPCSGCPOM 324
QY 283 DYILEVEPHPSSTKNLHLFQIMIR--KGTFAKYD-----YCLLGN 320
Db 325 GHYA--DRFSKRTDGVGQTFYLTNTDSSNFARWRYQVAVTSLGRRVTHVLVSLYSGSGN 382
QY 321 LRRY----GHLRPPAFDLSSIPESLPIWMYGGGLDALADVTDVQR-----TIR 364
Db 383 SKQYEIFTGLLKPQDTHLNE-----FSDVDVDGVDQVKVFWYNNVINPTLP 429

```

Oy 365 ELGSTPELLYICDYCHIDFVMSVRKAKDDVYVDL 397
::: : | : : | : |
Db 430 KVGASQITVEQNDGRVFKFCSTDTVREDILTL 462

Search completed: December 30, 2002, 20:44:30
Job time : 41 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2002, 20:09:51 ; Search time 92 Seconds
(without alignments)
918.253 Million cell updates/sec

Title: US-09-699-652A-14
Perfect score: 2194
Sequence: 1 MAMAGHAPGALPLILLWS.....DDVYVDFLRENGWHNSY 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phase:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-rviro:*
16: sp-bacteriap:*
17: sp-archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1055	48.1	344	10 Q9SKL5	Q9SKL5 arabidopsis
2	571.5	26.0	684	5 Q95XV1	Q95XV1 caenorhabdi
3	557.5	25.4	403	5 Q61866	Q61866 caenorhabdi
4	549.5	25.0	404	5 Q16936	Q16936 caenorhabdi
5	549	25.0	405	5 Q93789	Q93789 caenorhabdi
6	540	24.6	318	10 Q9FMT3	Q9FMT3 arabidopsis
7	539	24.6	395	11 Q9D798	Q9D798 mus musculu
8	539	24.6	395	11 Q9D611	Q9D611 mus musculu
9	538	24.5	395	11 Q9D7C5	Q9D7C5 mus musculu
10	538	24.5	395	11 Q9D6P3	Q9D6P3 mus musculu
11	537	24.5	395	11 Q9D767	Q9D767 mus musculu
12	537	24.5	395	11 Q9D6X0	Q9D6X0 mus musculu
13	537	24.5	395	11 Q9D6T5	Q9D6T5 mus musculu
14	537	24.5	395	11 Q9D6S5	Q9D6S5 mus musculu
15	537	24.5	395	11 Q9D6N8	Q9D6N8 mus musculu
16	537	24.5	395	11 Q9C9P7	Q9C9P7 mus musculu

17	536	24.4	395	11 Q9D606	Q9D606 mus musculu
18	536	24.4	395	11 Q9D6Q2	Q9D6Q2 mus musculu
19	535	24.4	395	11 Q9D796	Q9D796 mus musculu
20	535	24.4	395	11 Q9D619	Q9D619 mus musculu
21	531	24.2	395	11 Q9D6Q3	Q9D6Q3 mus musculu
22	530	24.2	395	11 Q9D766	Q9D766 mus musculu
23	530	24.2	395	11 Q9C9P8	Q9C9P8 mus musculu
24	529	24.1	395	11 Q9D6M9	Q9D6M9 mus musculu
25	528	24.1	395	11 Q9D760	Q9D760 mus musculu
26	502	22.9	374	4 Q96LG2	Q96LG2 homo sapien
27	500.5	22.8	426	5 Q17766	Q17766 caenorhabdi
28	498	22.7	456	5 Q9VKS5	Q9VKS5 drosophila
29	495.5	22.6	434	5 Q9VKT9	Q9VKT9 drosophila
30	474	21.6	411	5 Q20449	Q20449 caenorhabdi
31	470.5	21.4	351	5 Q95X33	Q95X33 caenorhabdi
32	468.5	21.4	399	5 Q9VPE9	Q9VPE9 drosophila
33	466.5	21.3	411	5 Q94252	Q94252 caenorhabdi
34	458	20.9	398	5 Q9V796	Q9V796 drosophila
35	425.5	19.4	838	5 Q9VBK6	Q9VBK6 drosophila
36	418	19.1	416	5 Q9VQ05	Q9VQ05 drosophila
37	417	19.0	457	5 Q9VKT2	Q9VKT2 drosophila
38	417	19.0	457	5 Q8T3X7	Q8T3X7 drosophila
39	416.5	19.0	559	5 Q17219	Q17219 bombyx mori
40	412.5	18.8	406	5 Q9VKT7	Q9VKT7 drosophila
41	410.5	18.7	616	5 Q77107	Q77107 plodia inte
42	401	18.3	504	5 Q94568	Q94568 galleria me
43	392.5	17.9	401	5 Q9U276	Q9U276 caenorhabdi
44	384.5	17.5	355	5 Q9VKT0	Q9VKT0 drosophila
45	379	17.3	435	5 Q9VG46	Q9VG46 drosophila

ALIGNMENTS

RESULT 1

ID	Q9SKL5	PRELIMINARY:	PRT:	344 AA.
AC	Q9SKL5;			
DT	01-MAY-2000 (TRENBLrel. 13, Created)			
DT	01-MAY-2000 (TRENBLrel. 13, Last sequence update)			
DT	01-MAR-2002 (TRENBLrel. 20, Last annotation update)			
DE	Putative lysosomal acid lipase.			
GN	AT2G15230.			
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
EX	MEDLINE=20083487; PubMed=10617197;			
RA	Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,			
RA	Fujii C.Y., Mason T.M., Bowman C.B., Barnstead M.E., Feldblum T.V.,			
RA	Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,			
RA	Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,			
RA	Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,			
RA	Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,			
RA	Salzberg S.L., Fraser C.M., Venter J.C.,			
RT	"Sequence and analysis of chromosome 2 of the plant Arabidopsis			
RT	thaliana.";			
RL	Nature 402:761-768(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RA	Lin X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AC006238; AAD25569.1; -			
DR	InterPro; IPR000734; Lipase.			
DR	InterPro; IPR000379; Ser_estrs_site.			
DR	PROSITE; PS0120; LIPASE_SER; 1.			
SEQ	SEQUENCE 344 AA; 38868 MW; 7BC4ED067E3602C4 CRC64;			

```
Query Match      48.1%; Score 1055; DB 10; Length 344;
Best Local Similarity 54.5%; Pred. No. 1.2e-88;
Matches 198; Conservative 42; Mismatches 71; Indels 52; Gaps 2;

Qy 40 GSGGGLCDQLLLPLGYPCTEHNHVTGDFLLSLQHHPHGKNAADSTGPPVFLQHLGFQ 99
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 26 GSPVNSLCADLIHPANYSCTEHSIQTKDGYILALQRVASLGPRL--QSGGPPVLLQHLGF 83

Qy 100 GGDWTFINSAQSLGYILADNGFDWIGNVGRGTRWSKGHSTFSDHKLFWDSHQELAEY 159
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 AGDVNLFNSLGSGLFILDHGFVWGNVGRGTRYSYGHVTLSDTDKEFDWDSWODLAMY 143

Qy 160 DLLAMLGVYVYVTSQKILYVGHSGGTIMGLAALTPEIVTKMISAAALCPISYLDHVSAS 219
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 144 DLAEIMQVLYSISKSLFVGHSGGTIMSPAALTQPHVAENVAAALCPISYLDHVTAP 203

Qy 220 FVLRAVAMHLDQMLVTMGHOLNFRSDMGVQIVDSLDCDGEHVDVNCNLLSATGNCFCNT 279
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 LVERVFMHLDQ----- 215

Qy 280 SRIDYILEYEPHPSPSTKNLHLLFOMIRKGTAKYDGLLGNLRYGHLRPPAFDLSSTPE 339
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 216 --IEYLDYEPHPSPSVKNIRLHFOMIRKGTFAQYDGYFKNLRYGSLKPPFILLSHIPA 273

Qy 340 SLPIWMGYGGLDALADVTQVIRELGSTPELLYIGDYGHIDFVMSYAKAKDDVYVDLIR 399
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 274 SLPMWVGYYGTDGLADVTQVIRELGSTPELLYIGDYGHIDFVMSYAKAKDDVYVDLIR 333

Qy 400 FRL 402
    || |
Db 334 FFR 336

RESULT 2
Q95XV1
ID Q95XV1 PRELIMINARY; PRT; 684 AA.
AC Q95XV1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 77.3 kDa protein.
GN Y57E12B.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RA Becker M.;
RT "The sequence of C. elegans cosmid Y57E12B.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024835; AAK68537.1;
DR InterPro; IPR00073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
KW Hypothetical protein.

SQ SEQUENCE 684 AA; 77339 MW; 85D3F46C9DE81F61 CRC64;
Query Match      26.08; Score 571.5; DB 5; Length 684;
Best Local Similarity 33.7%; Pred. No. 1.1e-43;
Matches 126; Conservative 75; Mismatches 144; Indels 29; Gaps 10;

Qy 49 QLLPLGLYPCTEHNHVTGDKGFLLSLQHHPHGKNAADSTG----PPVFLQHLGFQGGDTW 104
    : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 279 EIITHWGPVETHKVVTVGDIYLLHRIHPHGKNETSKSASKTPKPVFLQHLGLCTSSIW 338

Qy 105 FINSAESGLYILADNGFDWIGNVGRGTRWSKGHSTFSDHKLFWDSWQELAEYDLIAM 164
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 339 LLNLPQSGAGYIFADQGYDVLGNMGRNTYSKEHTRMTSADRRTWKESWEEMARYDLPA 398

Qy 165 LGY-VYVVTQSKILYVGHSGGTIMGLAALT-PEIVKMISSAAALCPISYLDHVSASFV- 221
    : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 399 INYALKTTKRONLYYVGHSGGALTMTFAKMSDEPMSKIRKFFAMAPVARMHSHVKGFLQN 458

Qy 222 LRAVAMHLDQMLVTMGHOLNFRSDMGVQIVDSLDCDGEHVD---CNLLSATIGENC-CF 277
    : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 459 LGOIYEQYNLYVQVFGDGEFLTNNFTKLLTDIFCD-QAVNPLCENFIFAVSGPNSNQF 517

Qy 278 NTSRIDYILEYEPHPSPSTKNLHLLFOMIRKGTAKYDGLLGNLRYGHLRPPAFDLSI 337
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 518 NNSRIGIYLAHPAGTSSRNILHFAQMVKKRMSRFDHGKDLNLKIYGAPSPPEYDIRKI 577

Qy 338 PESLPIMWGYGGLDALADVTQV-----RTIRELGSTPELLYIGDYGHIDFVMSVK 388
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 578 NSS--IYLFYSDFDLANPKDVEGLPMLPSKTLKKATK-----LRDFNHDFLWGM 629

Qy 389 AKDDVYVDLIRFLR 402
    | : | | | | | | | |
Db 630 ARKEIYDKIINTIK 643

RESULT 3
O61866
ID O61866 PRELIMINARY; PRT; 403 AA.
AC O61866
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 45.7 kDa protein.
GN ZK6.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RA Wu X.;
RT "The sequence of C. elegans cosmid ZK6.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067942; AAG45574.1;
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
KW Hypothetical protein.
```


Db	33	QIIMRWGYPAMIYDVNTEDG ¹ IIELRIRIPYGKTNVTPNGKKPVPVPMQHGLECCSSNVVV	92
QY	107	NSAEQSLGYILADNGFDVWGNVGRTRWSKGHSFTSVBHKLFWDMSWQELAEYDLMLAMLG	166
Db	93	NLPETSAFLFADAGYDVVLGNFRGNTYSMKHKNLKPSSAFWDSWDEMOQYDLPAMIE	152
QY	167	YVYTVT--QSKILYGVHSQGTINGLAALTAIPEI--VKMISSAALLCPSIYLDHVSAS--FV	221
Db	153	KALEVTGQDSLYYIGHSQGTLTFWSRLSEDKVGGWGNKIKFFALAPVGSVKHIKALAKFF	212
QY	222	LRAVAHMLDQMLVTMGIHQLNFRSDMGVQIVDSLCDGEHVD---CNNLLSAITG--ENCCF	277
Db	213	ADYFSLEFDDGDFVDFGSGEFLPNNTIMKLVSVCAGLKVEAGVCDVMFLTAGPENOL	272
QY	278	NTSRIDYILEYEPHPSTKNLHLFQWIRKGTFAKYDYGLIGLNLRRYGHLPAPADLSII	337
Db	273	NATRPVIYVAHTPAGTSTQNTVHWIQVRRHGTPKYDYGEKNKKHYGAQNPVADYFTTV	332
QY	338	PESLIPTWMGYGGDLADADTVTQIRELGSTPELL-----YIGDYGHIDFVM	385
Db	333	--NRPVYLYGDSDDLADPTDV-----TDFLLTLNPNSTVYQNNKLIDYNHLDFTW	381
QY	386	SVRAKDDVYVDLIRFLREN	404
Db	382	GLRAPKDIYPIIDIVRND	400

```

RESULT 6
Q9FWMT3 ID Q9FWMT3 PRELIMINARY; PRT; 318 AA.
AC Q9FWMT3;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE Similarity to lipase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RT Tabata S.;
RA "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
RD EMBL; AB007650; BAB08297.1; -.
SQ SEQUENCE 318 AA; 36448 MW; 25D72693E454743F CRC64;

```

	QY	314	DYGLLG-NLRRYCHLPAPFDLSISSIPESLPIWMVGYGLDALADVTDVQTIRELG-----S	368
	:	:	: : : :	:
	Db	222	NYGSORNIRIKHYCGAIPPAYINSAIPHLEFFSYGLDLSLDKVDVFLLDQFYHDID	281
	:	:	: : : :	:
	QY	369	TPELLXIGDXGHIDEVMYSKKAKDDVVYLRLRSL	402
	:	:	: : : :	:
	Db	282	KMNVPQVKDYAHADFIMGVTAKDGVYNQVATFEFK	315
	:	:	: : : :	:
	RESULT	7		
	ID	Q9D798	PRELIMINARY;	PRT; 395 AA.
	AC	Q9D798;		
	DT	01-JUN-2001	(TREMBlrel. 17, Created)	
	DT	01-JUN-2001	(TREMBlrel. 17, Last sequence update)	
	DT	01-MAR-2002	(TREMBlrel. 20, Last annotation update)	
	DE	231005I1B2LRik	protein..	
	GN	231005I1B2LRik.		
	OS	Mus musculus (Mouse).		
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
	NX	NCBI Taxid=10090;		

[illegible]

```

DR PROSIE, F300120, LIFASESER, I.
SQ SEQUENCE 395 AA; 44603 MW; D3FD8B6FEA671E3E CRC64;

Query Match 24.6%; Score 539; DB 11; Length 395;
Best Local Similarity 33.6%; Pred. No. 4 6e-41;
Matches 129; Conservative 68; Mismatches 135; Indels 52; Gaps

QY 49 QLLPLGYPCTEHNVTKGFLLSLQHPHGKKAADSTG--PPVFLQHLFGGGDTWFI 106
: : : : : : : : : : : : : : : : : : : : : :
Db 36 QMITYWCYPSEEVVTDGYILGVYRIPYGK-KSENIGKRPVAYLQHLGIASATNWIT 94
: : : : : : : : : : : : : : : : : : : : : :

QY 107 NSAEQSLGYILADNGFDVWITGVNRTGWSKGHSFTFSVHDKLFWDWSQELAEVLLAMLG 166
: : : : : : : : : : : : : : : : : : : : : :
Db 95 NLPNNSLAFILADAGYDVLWNLGRGNTWSRKNNYVSPDSVEEFWAFSFEDEMAKYDLPATID 154
: : : : : : : : : : : : : : : : : : : : : :

QY 167 VYVTVT--QSKILYVGHSGQITMGLAAL-TMPEIVTKMLSSAALLCPISYLDHVSAA----- 218
: : : : : : : : : : : : : : : : : : : : : :
Db 155 FIVOKTQGEKIHYVGHSGQTTIGFIAFSTNPALAKKIKRFPYALAPVATVKYTESPPKKIS 214
: : : : : : : : : : : : : : : : : : : : : :

QY 219 ---SFVLRAV-----AMHLDQMLVTMGIHOLFNRSDMGVQIVDSLCDGEHVD---CNN 265
: : : : : : : : : : : : : : : : : : : : : :

```

```

Db 215 LIPKFLKLVIFGNKMPHNYLDQFLGT-----EVCSELLDLCSN 256
QY 266 LLSAITG-ENCCFNTSRIDYILEYEPHPSSTKNLHFLFOMIRKGTFAKYDYL-LGNLRR 323
Db 257 ALFIFCGFDKKNLNSRLDVLGNHNPAGTSTQDLFWHAQLAKSGKLOAYNWGSPQLNMLH 316
QY 324 YGHLRPPAFDLSIPESLPIMWGYGLDALADVDQRTIRELGSTPELly---IGDYGH 380
Db 317 YNOKTPPYDVSAMTVPIAVW--NGGHDILADPQDVAMLLPKL---PNLLYHKELIPYNH 371
QY 381 IDFMVSVKAKDDVYVDLIRFLREN 404
Db 372 LDFIAMDAPQEVYNEIVTMAED 395

RESULT 8
Q9D6L1
ID Q9D6L1 PRELIMINARY; PRT; 395 AA.
AC Q9D6L1
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
clone:2310079020, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK010231; BAB26784.1;
DR MGD: MGI:1914567; 2310051B21RIK.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR pfam: PF00561; abhydrolase; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44659 MW; 1350BB95FC5A40E0 CRC64;

Query Match 24.6%; Score 539; DB 11; Length 395;
Best Local Similarity 33.9%; Pred. No. 4.6e-41;
Matches 130; Conservative 66; Mismatches 136; Indels 52; Gaps 13;

QY 49 QLLPLGYPCTEHNVTGDKGFLLSLQHIPGKNKAADSTG--PPVFLQHLFGGDTWFI 106
Db 36 QMITWYGPSEYEVVETDGYILGVYRIPYK-KNSENIGKRPVAYLQHLIASATNWT 94
QY 107 NSAEQSLGILLADGFDVWIGNVGRTRWSKGSTFVSHDKLFWDSWQBLAEYDLLAMLG 166
Db 95 NLPNLSAFILADAGYVWLGNSRGNTWSRKNVYSPDSVEFWAFSDEMAYIDLPATID 154

```

```

QY 167 YVYTVT-QSKILYVGHSGQTIMGLAAL-TWPEIVKVMISSAALLCPISYLDHVS----- 218
Db 155 FIVQKTGQEKIHYVGHSGQTTIGTAFSTNPALAKKIKRFYALAPVATVKYTESPEKKIS 214
QY 219 ---SFVLRV-----AMHLDQMLVTWGIHQINFRSDMGVOIVSLDCGHEVD---CNN 265
Db 215 LIHKFLKLVIFGNKMPHNYLDQFLGT-----EVCSELLDLCSN 256
QY 266 LLSAITG-ENCCFNTSRIDYILEYEPHPSSTKNLHFLFOMIRKGTFAKYDYL-LGNLRR 323
Db 257 ALFIFCGFDKKNLNSRLDVLGNHNPAGTSTQDLFWHAQLAKSGKLOAYNWGSPQLNMLH 316
QY 324 YGHLRPPAFDLSIPESLPIMWGYGLDALADVDQRTIRELGSTPELly---IGDYGH 380
Db 317 YNOKTPPYDVSAMTVPIAVW--NGGHDILADPQDVAMLLPKL---PNLLYHKELIPYNH 371
QY 381 IDFMVSVKAKDDVYVDLIRFLREN 404
Db 372 LDFIAMDAPQEVYNEIVTMAED 395

RESULT 9
Q9D7C5
ID Q9D7C5 PRELIMINARY; PRT; 395 AA.
AC Q9D7C5
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 2310051B21RIK protein.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK009359; BAB26240.1;
DR MGD: MGI:1914567; 2310051B21RIK.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR pfam: PF00561; abhydrolase; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44748 MW; 3ED43992458DE058 CRC64;

Query Match 24.5%; Score 538; DB 11; Length 395;
Best Local Similarity 34.1%; Pred. No. 5.7e-41;
Matches 131; Conservative 65; Mismatches 136; Indels 52; Gaps 14;

QY 49 QLLPLGYPCTEHNVTGDKGFLLSLQHIPGKNKAADSTG--PPVFLQHLFGGDTWFI 106

```


DR	MGI:1914967; 2310051B2IRik.
DR	InterPro; IPR000073; Abhydrolase.
DR	InterPro; IPR000734; Lipase.
DR	InterPro; IPR000379; Ser_estrs_site.
DR	Pfam; PF00561; abhydrolase_1.
DR	PROSITE; PS00120; LIPASE_SER; 1.
SQ	SEQUENCE 395 AA; 44607 MW; E43317C2254FA8FB CRC64;

Query Match		24.5%; Score 537; DB 11; Length 395;
Best Local Similarity		33.6%; Pred. No. 7e-41;
Matches 129; Conservative	67; Mismatches 136; Indels 52; Gaps 13;	

QY	49 QLLPLGLGPGCTEHNVETKDGFLLSLQHHPHGKNKAADSTG--PPVFLQHLGFQGGDTWFI 106
	: : : : : : : : : : : : : : : : : : : : :
DB	36 QMITYWPGESEYEVVTEDGYILGVYRYPYGK-KNSENIGKRVPAYLQHLGIASATNMIT 94
	: : : : : : : : : : : : : : : : : : : : :
QY	107 NSAEQSGLYLADNGFDWIGNVRGTWSKGHSFTSFVDHKLFWDSMOELAEYDLLAMLG 166
	: : : : : : : : : : : : : : : : : : :
DB	95 NLPNNSLAFLADAGYDVLGNSRGNTWSRKNNYYSPDSVEFWAFSDEMAKYDLDPATID 154
	: : : : : : : : : : : : : : : : : : :
QY	167 YVYPTVT-QSKILYVGHSQGTIMGUAAAL-TMPEIVKMISSAALLCPISYLDHVSA----- 218
	: : : : : : : : : : : : : : : : : : : : :
DB	155 FIVQKTGOEKIHYVGHSGTGAIGFIAFSTNPALAKKIKRFVALAPVATVKYTESPKKIS 214
	: : : : : : : : : : : : : : : : : : : : :
QY	219 ---SFVLRAV-----AHILDONLYMTGHQLNFSDMGVQIVDSLDCGEHVD--CNN 265
	: : : : : : : : : : : : : : : : : : : : :
DB	215 LIPAKLFKVIFGNKMFPHNYLDQFLGT-----EYCSRELLDLLCSN 256
	: : : : : : : : : : : : : : : : : : : : :
QY	266 LLSAITG-ENCCENTSRDIYLEYEPHPSSPKNLHHLFQMIRKGFKAQDYGL-LGNLRR 323
	: : : : : : : : : : : : : : : : : : :
DB	257 ALTFCEGDKNLNVSREDYVLGHNPAGTSQDLPFHQAQKSLQAINWGSPLONNMLH 316
	: : : : : : : : : : : : : : : : : : :
QY	324 YGHURPPAFLSSIPESLPIWMGYSGLDALADVQRTIRELGSTPELLY---IGDYGH 380
	: : : : : : : : : : : : : : : : : : :
DB	317 YNQKTPPYDVSAVMTPVIAVN--NGHDILADPDQVAMLLPKL---PNLLYHKELLPLYNH 371
	: : : : : : : : : : : : : : : : : : :
QY	381 IDFMVSVKAKDDYVDLIRFLREN 404
	: : : : : : : : : : : : : : : : : : :
DB	372 LDFTWAMDAPQEYNEIVTMMAED 395
	: : : : : : : : : : : : : : : : : : :

RESULT 12	
Q9DXO	PRELIMITINARY; PRT; 395 AA.
ID	Q9DXO PRELIMITINARY; PRT; 395 AA.
AC	Q9DXO
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBurel. 17, Last sequence update)
DT	01-MAR-2002 (TrEMBurel. 20, Last annotation update)
DE	2310051B2IRik protein.
GN	2310051B2IRik.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=TONGUE;
RX	MEDLINE=21085660; PubMed=11217851;
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA	kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA	Schral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA	L Lyons P., Marchionni L., Mashima J., Mezzarelli J., Mommaerts P.,
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

```

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009990; BAB26229.1; -.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; D3F96DB83161C3EF CRC64;

Query Match 24.5%; Score 537; DB 11; Length 395;
Best Local Similarity 33.6%; Pred. No. 7e-41;
Matches 129; Conservative 67; Mismatches 136; Indels 52; Gaps 13;

Qy 49 QLLPLGYPCTEHNVTGDKGFLLSLQHIPGKNAADSTG--PPVFLOHGLFQGGDTWFI 106
Db 36 QMITWGYPSEYEYVETDGYILGVYRIPYCK-KNSENIGKRPVAYLQHGGLIASATNMIT 94

Qy 107 NSAEQSLGYILADGDFWVGNVGRTRWSKGHSFTSVHDKLFWDSWQELAEYDILLMLG 166
Db 95 NLPNNSLAFLADAGYDVLGNSRGNTWSRKNVYSPDSVEFWAFSDEMAYKYLDPATID 154

Qy 167 VYVTVT-OSKILYVGHSGGTTMGLAAL-TMPEIVKMISAAALCPISYLDHVSA----- 218
Db 155 FIVQKTGGEKIHVGHSGGTTIGFTAFSTNPAKAKIKRFYALAPVATVVKYTESPFKKIS 214

Qy 219 ---SEFLRAV-----AMHLDQMLVTMGTHQLNFRSDMGVQIVDSLCDGEHVD--CNN 265
Db 215 LIPKFLKLVIFGNKMFPHNYLDQFLGT-----EVCSSRELLDLCNS 256

Qy 266 LLSAITG-ENCCFNTSRIDYILEYEPHPSSTKNLHLFQMIKRGTFKAYDYGL-LGNLRR 323
Db 257 ALFIFCGFDKKNLVSREFDVLGHNPACTSTQDLFWAQLAKSGKLQAYNNGSPQLNNMLH 316

Qy 324 YGHLRPPAFDLSSIPESLPINWGYGLDALADVTQVIRTRELSTPELly---IGDYGH 380
Db 317 YNOKSPPIYDVSAMTVPIAVW--NGGHDILADPDQVAMLLPKL---PNLLYHKEILPYNH 371

Qy 381 IDFVMSVKAKDDVYVDLIRFLREN 404
Db 372 LDFIWAMDAPQEVYNEIVTMAED 395

RESULT 14
Q9D6S5 PRELIMINARY; PRT; 395 AA.
AC Q9D6S5;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310063K07, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

```

```

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W.J., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010026; BAB26651.1; -.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44650 MW; 03F16D53373A4D57 CRC64;

Query Match 24.5%; Score 537; DB 11; Length 395;
Best Local Similarity 33.6%; Pred. No. 7e-41;
Matches 129; Conservative 67; Mismatches 136; Indels 52; Gaps 13;

Qy 49 QLLPLGYPCTEHNVTGDKGFLLSLQHIPGKNAADSTG--PPVFLOHGLFQGGDTWFI 106
Db 36 QMITWGYPSEYEYVETDGYILGVYRIPYCK-KNSENIGKRPVAYLQHGGLIASATNMIT 94

Qy 107 NSAEQSLGYILADGDFWVGNVGRTRWSKGHSFTSVHDKLFWDSWQELAEYDILLMLG 166
Db 95 NLPNNSLAFLADAGYDVLGNSRGNTWSRKNVYSPDSVEFWAFSDEMAYKYLDPATID 154

Qy 167 VYVTVT-OSKILYVGHSGGTTMGLAAL-TMPEIVKMISAAALCPISYLDHVSA----- 218
Db 155 FIVQKTGGEKIHVGHSGGTTIGFTAFSTNPAKAKIKRFYALAPVATVVKYTESPFKKIS 214

Qy 219 ---SEFLRAV-----AMHLDQMLVTMGTHQLNFRSDMGVQIVDSLCDGEHVD--CNN 265
Db 215 LIXKFLKLVIFGNKMFPHNYLDQFLGT-----EVCSSRELLDLCNS 256

Qy 266 LLSAITG-ENCCFNTSRIDYILEYEPHPSSTKNLHLFQMIKRGTFKAYDYGL-LGNLRR 323
Db 257 ALFIFCGFDKKNLVSREFDVLGHNPACTSTQDLFWAQLAKSGKLQAYNNGSPQLNNMLH 316

Qy 324 YGHLRPPAFDLSSIPESLPINWGYGLDALADVTQVIRTRELSTPELly---IGDYGH 380
Db 317 YNOKTPPIYDVSAMTVPIAVW--NGGHDILADPDQVAMLLPKL---PNLLYHKEILPYNH 371

Qy 381 IDFVMSVKAKDDVYVDLIRFLREN 404
Db 372 LDFIWAMDAPQEVYNEIVTMAED 395

RESULT 15
Q9D6N8 PRELIMINARY; PRT; 395 AA.
AC Q9D6N8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310074E21, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2002, 20:42:42 ; Search time 37 seconds
(without alignments)
326.038 Million cell updates/sec

Title: US-09-699-652A-14
Perfect score: 2194
Sequence: 1 MAMAGHAPGALPLILLVWS.....DDVYVDLIRFLRENGWHSY 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/1aa/PC10S_COMB.pap.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	571	26.0	378	4	US-09-186-489-2
2	535.5	24.4	380	1	US-08-227-108-16
3	535.5	24.4	380	2	US-09-073-674-16
4	534	24.3	377	2	US-08-227-108-17
5	534	24.3	377	2	US-09-073-674-17
6	529	24.1	379	1	US-08-227-108-3
7	529	24.1	379	2	US-09-073-674-3
8	529	24.1	380	1	US-08-227-108-5
9	529	24.1	380	2	US-09-073-674-5
10	507.5	23.1	392	4	US-09-820-001-4
11	492.5	22.4	395	4	US-09-820-001-2
12	491.2	22.4	379	1	US-08-227-108-18
13	491	22.4	379	2	US-09-073-674-18
14	105	4.8	5588	4	US-09-036-987A-6
15	105	4.8	5588	4	US-09-370-700-6
16	94	4.3	634	1	US-07-779-049-3
17	94	4.3	634	1	US-08-080-240-3
18	93.5	4.3	1167	1	US-08-100-709-2
19	93.5	4.3	1167	1	US-08-176-865-2
20	93.5	4.3	1167	1	US-08-474-038-2
21	93.5	4.3	1167	2	US-08-779-046-2
22	93.5	4.3	1167	2	US-08-881-340-2
23	93	4.2	388	1	US-08-232-519-2
24	93	4.2	388	1	US-08-456-956-2
25	90.5	4.1	465	4	US-08-985-432-15
26	90	4.1	1676	4	US-08-487-283A-2
27	88	4.0	972	3	US-08-750-141A-2

28	87.5	4.0	298	4	US-09-355-166-6	Sequence 6, Appli
29	86	3.9	686	1	US-08-350-884-70	Sequence 70, Appl
30	86	3.9	686	1	US-08-709-173-70	Sequence 70, Appl
31	86	3.9	686	2	US-08-709-177-70	Sequence 70, Appl
32	86	3.9	841	1	US-08-350-884-86	Sequence 86, Appl
33	86	3.9	841	1	US-08-709-173-86	Sequence 86, Appl
34	86	3.9	841	2	US-08-709-177-86	Sequence 86, Appl
35	85	3.9	385	3	US-09-053-866-2	Sequence 2, Appli
36	85	3.9	385	4	US-09-479-130-2	Sequence 2, Appli
37	85	3.9	878	4	US-09-087-277-2	Sequence 2, Appli
38	85	3.9	878	4	US-09-658-499-2	Sequence 2, Appli
39	85	3.9	906	4	US-09-134-001C-3218	Sequence 3218, Ap
40	84.5	3.9	566	2	US-08-920-095-3	Sequence 3, Appli
41	84.5	3.9	566	5	PCT-US96-05800-3	Sequence 26, Appl
42	84.5	3.9	569	2	US-08-467-822-26	Sequence 26, Appl
43	84.5	3.9	569	4	US-08-432-697-26	Sequence 26, Appl
44	84.5	3.9	569	4	US-08-466-248-26	Sequence 26, Appl
45	84.5	3.9	1227	1	US-08-448-170-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-186-489-2
; Sequence 2, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; TITLE OF INVENTION: Purified Recombinant Kid pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Kid (Goat)
US-09-186-489-2

Query Match	26.0%	Score 571;	DB 4;	Length 378;
Best Local Similarity	35.5%	Pred. No. 1.le-53;		
Matches 130;	Conservative 66;	Mismatches 154;	Indels 16;	Gaps 9;
Qy	49	QLLLPLGYPCTEHNVTEDGFLLSQHIPGKNKAAD-STGPPVFLQHLFGGGDTWFIN	107	
Db	17	QMISFWGYPSEMHKVTADGYILQVYRIPHGKNDANHLGQRPVFLQHLGLASATNWSN	76	
Qy	108	SAEQSLGVLADNGFDVJGNGVGRTRMSKGHSTFVSHDKLFWDSWQELAEYDLAMLG	167	
Db	77	LPNNSLGLLADAGYDVWLGNSRGNTWAQEHLYSPDSPEFWAFSFDENAEYDLSPIDF	136	
Qy	168	VYTVT-QSKILYVGHSGQITMGLAAL-TMPEIVKMISAAALCPISYLDHVSASFVLRV	225	
Db	137	ILKRTGQKLLHVTGHSQGTIGFVAFSTNLAKEIVFHALAPVATVKTOSLTKKAL	196	
Qy	226	AMHLDQMLV---TMGIHOLNFRSDMGVQITVDSLCDGEHVD---CNNLLSAITG-ENCCFNT	279	
Db	197	IPHFLFKIIFGNKMEYPHNFFEQFLGVEV---CSRETLDVLCKNALFAITCADKNFNW	252	
Qy	280	SRIDYLYEYEPHPSKTNLHLFQMRKGTFAKYDYL-LGNLRRYGHILRPPAFDLSSTP	338	
Db	253	SRLDYVAHNPAQASVQNILHWRQAISKGFQAFDQWASVENLMHYNQPTPIYNLTAMN	312	
Qy	339	ESLPTWNGYGGDLADYTDVQRTIRELGSTPELILYIGDYGHIDFMSVKAKDDVYVDLI	398	
Db	313	VFIAYWS--AGODLLADPDQVDLLLSKLSNLTHHKEIPNYNHLDEFTWADAPQEVYNEII	370	
Qy	399	REFLREN 404		

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

```

Qy 219 ---SFVLRAVAMH-----LQDMLVTMGIHOLNFRSDMGVOIVDSLCDGEHVD---CNN 263
   ||: |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 LVPSFEKFLIGNKIFPHPHFFDQFLAT-----EVGSRFTVDLLCSN 238
   ||: |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 266 LLSAITG-ENCCEWTSRIDYLYVEVPHSPSTKNLHLFQMKRGTFKAYDYL-LGNLRR 323
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 239 ALFTICGFTWNUNMSRLDVLSHNPAGTSVQNVLHWSQAVKSGKQAFDMSGSPVQNMHH 298
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 324 YGHLRPPAFDLSSTPESLPIMWGGGLDALADYTDVQRTIRELGSTPELly---IGDYGH 380
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 299 YHQSMPPYYNLTDHVPPIAWN--NGGNDLLADPHDVL LLSKL---PNLIYHRKIPPNH 353
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 381 IDFMVSVKAKDDVYVDLIREL 401
   ||: |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 354 LDFIAMDAPQAVNYEIVSMN 374
   ||: |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-073-674-3
; Sequence 3, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-3

```

[illegible]

```

167 YVTVT--QSKILYVGSQGTIMGAL--TWPEIVKVMISSAALCPTISYLDHVS----- 21b
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 FILKKTQDQLKLVHVGSHQGTGIFIAFSTNPKLAKRIKTFYALAPVATVKYTTLLNKL 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 ---SEVLRAVAMH-----LDQMLVTMGIHOLNFRSDMGVQIVDSLCDGHEVD--CNN 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
198 LVFSFLKFLFGNKFIYPHFHFFQFLAT-----EVCSTRETVDLJCSN 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
266 LLSAITG-ENCCFNTRSKIDYIEYPHPSSTKNLHLFQMIRKGTFAKYDYL-LGNLRR 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 ALFTICGFDFTMNLNMSRLDYLSHNPAGTSVQNVLHWSQAVKSGKFOAFDWGSPVQNM 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 YGHLRPPAFDLSPTSPESLPTWVGYGGLDALADVTDTQRTIRELGSTPELly---ICDYGH 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 YHOSMPYYNLTMHVPDIAVW--NGGNDLLADPHVDLLSKL---PNLIYHRKTIPPYNH 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
381 IDFMVMSVKAKDDVYVDLIRFL 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 LDFIWMADPAQVYNEIVSMW 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-073-674-5
: Sequence 5, Application US/09073674
: Patent No. 5998189
: GENERAL INFORMATION:
: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Warner-Lambert Company
: STREET: 2800 Plymouth Road
: CITY: Ann Arbor
: STATE: Michigan
: COUNTRY: U.S.A.
: ZIP: 48105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/073,674
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Crissey, Todd M.
: REGISTRATION NUMBER: 37,807
: REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 734 622-7530
: TELEFAX: 734 622-1553
: TELEX:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 380 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-073-674-5

```


STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-18

Query Match 22.4%; Score 491; DB 1; Length 379;
Best Local Similarity 31.7%; Pred. No. 6.2e-45;
Matches 117; Conservative 78; Mismatches 152; Indels 22; Gaps 12;

QY 49 QLLPLGYPCTHHVETKDGFLLSLOHPIHG-KNKAADSTGPPVFLQHLFGGDTWFIN 107
Db 18 QMISWGYPSEKYEYVETDGYILEVNRIPYKKNNGRQRPVFLQHLGLASASNNWISN 77
QY 108 SAEQSLGILADNGFDVWIGNVGRTRWSKGHSTFVSHDKLFWDSWQELAEYDILLAMLG 167
Db 78 LPNNSLAFLADAGYGVWLGNSRGNTWSRNLYYSPDSVEFWAFSDEMAKYDLPATIDF 137
QY 167 VYTVT-QSKILYVGHSGQITMGLAAL-TMPEIVKMISSAALLCPISYLDHVSASFVLRV 225
Db 138 IVKETGQEKLVHVGHSQGTIGFIAFSTNPKLAERIKTFYALAPVATVKY-TKSLVNLK 196
QY 226 AMHLDQMLVTMG---IHOLNFRSDMGVQIVDSLCDGE--HVCNLLSAITG-ENCCFNT 279
Db 197 FIPPTMFKIIFGDKIFYPHNFFDQF---LATQVCSRETINLVICSNALFTICGDSANLNM 253
QY 280 SRIDYLYEPEHPSPSTKNLHLLFQMRKGTFAKYDGL-LGNLRRYGHLPAPFDLSSTP 338
Db 339 ESLPTMVGGLDALADVTQVORTIRELGSTPELly---IGDYGHIDFVMSVKAKDDYV 395
Db 314 VPIAVMS--GGNDWLADPOVDVLLPKLSN---LIYHKEILPYNHLDFTWAMNAPQEVYN 368
QY 396 DLIRFLREN 404
Db 369 EIIISMWAKD 377

RESULT 13
US-09-073-674-18
Sequence 18, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-18

Query Match 22.4%; Score 491; DB 2; Length 379;
Best Local Similarity 31.7%; Pred. No. 6.2e-45;
Matches 117; Conservative 78; Mismatches 152; Indels 22; Gaps 12;

QY 49 QLLPLGYPCTHHVETKDGFLLSLOHPIHG-KNKAADSTGPPVFLQHLFGGDTWFIN 107
Db 18 QMISWGYPSEKYEYVETDGYILEVNRIPYKKNNGRQRPVFLQHLGLASASNNWISN 77
QY 108 SAEQSLGILADNGFDVWIGNVGRTRWSKGHSTFVSHDKLFWDSWQELAEYDILLAMLG 167
Db 78 LPNNSLAFLADAGYGVWLGNSRGNTWSRNLYYSPDSVEFWAFSDEMAKYDLPATIDF 137
QY 167 VYTVT-QSKILYVGHSGQITMGLAAL-TMPEIVKMISSAALLCPISYLDHVSASFVLRV 225
Db 138 IVKETGQEKLVHVGHSQGTIGFIAFSTNPKLAERIKTFYALAPVATVKY-TKSLVNLK 196
QY 226 AMHLDQMLVTMG---IHOLNFRSDMGVQIVDSLCDGE--HVCNLLSAITG-ENCCFNT 279
Db 197 FIPPTMFKIIFGDKIFYPHNFFDQF---LATQVCSRETINLVICSNALFTICGDSANLNM 253
QY 280 SRIDYLYEPEHPSPSTKNLHLLFQMRKGTFAKYDGL-LGNLRRYGHLPAPFDLSSTP 338
Db 254 SRLDYVSHNPAGTSVQNNLHWTQAVKSGNFOAFNMGSPAQNVVHFNOPTPPYVNVVTAMN 313
QY 339 ESLPTMVGGLDALADVTQVORTIRELGSTPELly---IGDYGHIDFVMSVKAKDDYV 395
Db 314 VPIAVMS--GGNDWLADPOVDVLLPKLSN---LIYHKEILPYNHLDFTWAMNAPQEVYN 368
QY 396 DLIRFLREN 404
Db 369 EIIISMWAKD 377

RESULT 14
US-09-036-987A-6
Sequence 6, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.

APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-6

Query Match 4.8%; Score 105; DB 4; Length 5588;
Best Local Similarity 21.6%; Pred. No. 0.75;
Matches 87; Conservative 62; Mismatches 164; Indels 90; Gaps 19;
Qy 11 ALPLILLVSCCGRTVSCASPA-----AALRVG-----SGSGGLCDQL 50
Db 2232 AVPL-----IVSGRTPAALSAQASALLSYLGERGDI STLDAAFSLASSRAALEER 2281
Qy 51 LLPLGYPCTEHNVTETKDGFLLSLQHIPHGKNAADSTGPPV-----FLQHGLFGGDTWF 105
Db 2282 AVVLG-----ADRETLISGLEALASGREASGVSGSPVSGGVGFVAG--QGQWL 2331
Qy 106 INSAEQSLGYILAD-----NGFDVWIGNVRGTRWSKGHSTFVHDKLFDWDSWQELAEY 159
Db 2332 MGRGLYSVPVFADAFDEACAGLDAHLGQDVGVDRDVFSGSLDRTLWAQSGLFALOV 216
Qy 160 DLLAMLGYYVTQSKILYVGHSGQTINGLAA---LTMPETVKMTSSAALLCPSIYLDHV 216
Db 2332 MGRGLYSVPVFADAFDEACAGLDAHLGQDVGVDRDVFSGSLDRTLWAQSGLFALOV 216
Qy 160 DLLAMLGYYVTQSKILYVGHSGQTINGLAA---LTMPETVKMTSSAALLCPSIYLDHV 216
Db 2332 GLLSLLG--SWGVRPGVVL--GHSVGEFAAAVAGVLSLPDAARVAGRARL-----MOAL 2443
Qy 217 SASFVLRAVAMHLDQMLVTM-----GIHQLNFRSDMGVQIVDSLCDGHEHVCNLLSA 269
Db 2444 PSGGAMLAVAAAGEEQLRPLADRVGDGAGIAAVNAPESV-----VLSGDREVLDIAGA 2496
Qy 270 ITGENCCFNTSRIDY-YLEYEPHPSSTKNLHLLFQMIKRGTFKAYDY---GLLGNLRRY 324
Db 2497 LDGQGIKRRRLRVSHAFHSYRMDPM---LOEFAEIAIRSVDRRGDLPVSVTLTGELDTA 2552
Qy 325 GHLRPPAFDLSSIPESLPIMWGYGGLDALADVTVDQRTIRELG 367
Db 2553 GVMATPEYVWVROVRE--PVREA-DGVRVLAQ--QGVATIFELG 2590

Search completed: December 30, 2002, 20:48:12
Job time : 42 secs

RESULT 15
US-09-370-700-6
Sequence 6, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 5588
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

Query Match 4.8%; Score 105; DB 4; Length 5588;
Best Local Similarity 21.6%; Pred. No. 0.75;
Matches 87; Conservative 62; Mismatches 164; Indels 90; Gaps 19;
Qy 11 ALPLILLVSCCGRTVSCASPA-----AALRVG-----SGSGGLCDQL 50
Db 2232 AVPL-----IVSGRTPAALSAQASALLSYLGERGDI STLDAAFSLASSRAALEER 2281
Qy 51 LLPLGYPCTEHNVTETKDGFLLSLQHIPHGKNAADSTGPPV-----FLQHGLFGGDTWF 105
Db 2282 AVVLG-----ADRETLISGLEALASGREASGVSGSPVSGGVGFVAG--QGQWL 2331
Qy 106 INSAEQSLGYILAD-----NGFDVWIGNVRGTRWSKGHSTFVHDKLFDWDSWQELAEY 159
Db 2332 MGRGLYSVPVFADAFDEACAGLDAHLGQDVGVDRDVFSGSLDRTLWAQSGLFALOV 216
Qy 160 DLLAMLGYYVTQSKILYVGHSGQTINGLAA---LTMPETVKMTSSAALLCPSIYLDHV 216
Db 2332 GLLSLLG--SWGVRPGVVL--GHSVGEFAAAVAGVLSLPDAARVAGRARL-----MOAL 2443
Qy 217 SASFVLRAVAMHLDQMLVTM-----GIHQLNFRSDMGVQIVDSLCDGHEHVCNLLSA 269
Db 2444 PSGGAMLAVAAAGEEQLRPLADRVGDGAGIAAVNAPESV-----VLSGDREVLDIAGA 2496
Qy 270 ITGENCCFNTSRIDY-YLEYEPHPSSTKNLHLLFQMIKRGTFKAYDY---GLLGNLRRY 324
Db 2497 LDGQGIKRRRLRVSHAFHSYRMDPM---LOEFAEIAIRSVDRRGDLPVSVTLTGELDTA 2552
Qy 325 GHLRPPAFDLSSIPESLPIMWGYGGLDALADVTVDQRTIRELG 367
Db 2553 GVMATPEYVWVROVRE--PVREA-DGVRVLAQ--QGVATIFELG 2590

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw_model

Run on: December 30, 2002, 20:43:42 ; Search time 71 Seconds
(without alignments)
99.983 Million cell updates/sec

Title: US-09-699-652A-14
Perfect score: 2194
Sequence: 1 MAMAGHAPGALPILLVVS.....DDVYVDLIRFLRENGHNSY 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 109717 seqs, 17314136 residues

Total number of hits satisfying chosen parameters: 109717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	546	24.9	398	9	US-10-042-431-75
2	531	24.2	398	9	US-10-056-744B-2
3	515.5	23.5	399	9	US-10-042-431-41
4	507.5	23.1	392	12	US-10-003-302-4
5	502	22.9	374	10	US-09-811-825-4
6	502	22.9	390	9	US-10-042-431-49
7	502	22.9	398	10	US-09-811-825-2
8	502	22.9	409	10	US-09-835-996A-21
9	502	22.9	423	9	US-10-042-431-47
10	492.5	22.4	395	12	US-10-003-302-2
11	384.5	17.5	221	9	US-10-042-431-50
12	264	12.0	127	9	US-10-056-744B-5
13	114.5	5.2	144	9	US-10-042-431-52
14	90.5	4.1	465	10	US-09-923-779-151
15	90.5	4.1	473	10	US-09-925-297-567
16	90	4.1	1602	10	US-09-778-927A-59
17	89	4.1	305	10	US-09-815-242-12337
18	88	4.0	972	10	US-09-944-807-10
19	87.5	4.0	298	10	US-09-950-368-6

20	87.5	4.0	1251	10	US-09-778-927A-58	Sequence 58, Appl
21	87	4.0	315	12	US-10-080-644-2	Sequence 2, Appl
22	85.5	3.9	297	12	US-10-080-644-7	Sequence 7, Appl
23	85.5	3.9	2799	9	US-10-151-736-4	Sequence 4, Appl
24	85	3.9	232	10	US-09-896-578-4	Sequence 4, Appl
25	84.5	3.9	569	9	US-09-895-913A-252	Sequence 252, App
26	84	3.8	1079	9	US-09-895-793-947	Sequence 947, App
27	84	3.8	1079	10	US-09-822-827-947	Sequence 947, App
28	83.5	3.8	252	10	US-09-452-239-14	Sequence 14, Appl
29	83.5	3.8	467	10	US-09-735-933-6	Sequence 6, Appl
30	82.5	3.8	676	10	US-09-815-242-13513	Sequence 13513, A
31	82	3.7	686	9	US-09-881-239-3	Sequence 3, Appl
32	82	3.7	686	10	US-09-881-654-2	Sequence 2, Appl
33	82	3.7	728	9	US-09-881-239-1	Sequence 1, Appl
34	81	3.7	632	10	US-09-929-955-29	Sequence 29, Appl
35	81	3.7	686	10	US-09-929-955-17	Sequence 17, Appl
36	81	3.7	686	10	US-09-929-955-31	Sequence 31, Appl
37	81	3.7	686	10	US-09-929-955-32	Sequence 32, Appl
38	81	3.7	686	10	US-09-929-955-43	Sequence 43, Appl
39	81	3.7	686	10	US-09-929-955-44	Sequence 44, Appl
40	81	3.7	686	10	US-09-929-955-45	Sequence 45, Appl
41	81	3.7	686	10	US-09-929-955-46	Sequence 46, Appl
42	81	3.7	686	10	US-09-929-955-47	Sequence 47, Appl
43	81	3.7	686	10	US-09-929-955-48	Sequence 48, Appl
44	81	3.7	686	10	US-09-929-955-49	Sequence 49, Appl
45	80	3.6	2894	10	US-09-941-611-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-10-042-431-75
; Sequence 75, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-75

Query Match	24.9%	Score 546;	DB 9;	Length 398;
Best Local Similarity	33.9%	Pred. No. 2.1e-44;		
Matches 130;	Conservative 70;	Mismatches 132;	Indels 52;	Gaps 13;
QY	49	QLLLPLGYPTCHHNETKDGFLLSLQHPHGKNAADSTG--PPVFLQHLFGQGDWTWFI	106	
Db	37	QMITWGYPNEEYVETDGYILEVNRIPYGR-KNSGNTGQRPVFLQHLGLASATNWIS	95	
QY	107	NSAEOSLGVILADNGFDVWIGNVGRTRMSKGHSTSVHDKLFWDSWQELAYDILLAMLG	166	
Db	96	NLPNNSLAFILADAGYVGLWNLGRGNTWARRNLYTSPDSVEFWAFSEDMAKYDLPATID	155	
QY	167	VYVYTVT-QSKILYVGHSGQTINGLAAL-TMPEIVKMISSAALLCPIISLDHVSASF-VLR	223	
Db	156	FIVKTKGQKQLHYVGHSGQTTGTFATFSTPNPSLAKRIKTFYALAPVATVYKTKSLINKLR	215	

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-302-4

Query Match 23.1%; Score 507.5; DB 12; Length 392;
Best Local Similarity 32.4%; Pred. No. 1e-40;
Matches 130; Conservative 57; Mismatches 141; Indels 73; Gaps 12;

QY 42 GSGGL-----CDQLLPLGVPCTEHNVTGKGLLSLQHIHPG-KNKAADSTGP 89
Db 19 GSGGLTAVDPETNMNVSEIISYMGFPSEYLVETEDGYILCLNRPCHKRKHSDGPKP 78
QY 90 PVFLQHLFGQDGTWFINSAGOSGLVILADGFDVWGNVGRTRSKGHSFVHDKLEW 149
Db 79 VVFLQGLLADSSNWTNLANSLSGLFADAGFDVWGNRGTWSRKHKTLVSQDEFW 138
QY 150 DMSWQELAEYDLLAMLGYVYTVT-OSKILYVGHSGQTINGLAALT-MPEIVKMISSAALL 207
Db 139 AFSYDEMAKYDLPASINFILNLTQGOVYVYVGHSGQTTIGTAFASQIPELAKRIKMFAL 198
QY 208 CPISYLDHVSASFVLRAVAM-----HLDQMLVTMGIHOLNFRSDMGVQIVDSLCD----- 257
Db 199 GPVA-----SVAFCTSPMAKLGRLPDHL-----IKDLFGDKLEFL 232
QY 258 -----GEHVDCNNLSAITGENCEF-----NTSRIDYLEYEPHPSKML 298
Db 233 POSAFKWLGTHTVCHVILKECG-NLCELLCGFENRNLNMSRDVYVTHSPAGTSVQNM 291
QY 299 HHLFQMRKGTFAKYDYLGLL-NLRRYGHLPAPFDLSSIPESLPIWMYGGDLADAVT 357
Db 292 LHSQAVKFKQFQADGWSAKNYHYNOSYPTYNKDLVPTAVWS--GGHDLADYI 349
QY 358 DVORTIRELGSTPELLYIGDYHIDFVMSVKAKDDVYVDLI 398
Db 350 DNVILLTQITNLVFNESIPWEHDLDFIWLGDAPWRLYNKII 390

RESULT 5

US-09-811-825-4
; Sequence 4, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:

; APPLICANT: VAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001170
; CURRENT APPLICATION NUMBER: US/09/811,825
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Human
US-09-811-825-4

Query Match 22.9%; Score 502; DB 10; Length 374;
Best Local Similarity 34.0%; Pred. No. 3.2e-40;
Matches 124; Conservative 67; Mismatches 146; Indels 28; Gaps 13;

QY 55 GYPCTEHNVTGDKGFLLSLQHIHPGKNKAADSTG--PPVFLQHLFGQDGTWFINSAGOS 112
Db 8 GYPCEEYEVATEDGYILSVNRPGLVQ-PKKTGSRPVVLLQHLGVGGASNNISLNPNS 66
QY 113 LGYILADGFDVWGNVGRTRSKGHSFVHDKLFWDSWQELAEYDLLAMLGYVYTVT 172
Db 67 LGFILLADAGFDVWGNRGNASRKHKTLSDIDODEFAFSDYDEMARFDLPVINFILQKT 126

QY 173 -OSKILYVGHSGQTINGLAAL-TMPEIVKMISSAALLCPISYLDHVS-----SFVLRAVAM 227
Db 127 GOEKIYVYGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPIATVKHAKSPCTKFL----- 181
QY 228 HLDQMLV--TMGIHOLNFRSDMGVQIVDSLCDGEHVD--CNNLLSAITGENC-CFNTSRI 282
Db 182 -LPDMWIKGLFGKKEFLYQTRFLRQLVYLQGVILDQICSNIMLLLGGFTNNMMSRA 240
QY 283 DYLYEVEPHPSSTKNLHLLFQMRKGTFAKYDYGL-LGNLRRYGHLPAPFDLSSIPESL 341
Db 241 SYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVRYRVDM--TV 298
QY 342 PIWMYGGDLADAVTDVORTIRELGSTPELLY---IGDYGHIDFVMSVKAKDDVYVDLI 398
Db 299 PTAMTGGQDLSNPEDVKMLLSEV---TNLIYHKNIPEMAHVDFIWLGDAPHRMYNEII 355
QY 399 RFLRE 403
Db 356 HLMQO 360

RESULT 6

US-10-042-431-49
; Sequence 49, Application US/10042431
; Publication No. US20020182675A1

; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-49

Query Match 22.9%; Score 502; DB 9; Length 390;
Best Local Similarity 34.0%; Pred. No. 3.3e-40;
Matches 124; Conservative 67; Mismatches 146; Indels 28; Gaps 13;

QY 55 GYPCTEHNVTGDKGFLLSLQHIHPGKNKAADSTG--PPVFLQHLFGQDGTWFINSAGOS 112
Db 24 GYPCEEYEVATEDGYILSVNRPGLVQ-PKKTGSRPVVLLQHLGVGGASNNISLNPNS 82
QY 113 LGYILADGFDVWGNVGRTRSKGHSFVHDKLFWDSWQELAEYDLLAMLGYVYTVT 172
Db 83 LGFILLADAGFDVWGNRGNASRKHKTLSDIDODEFAFSDYDEMARFDLPVINFILQKT 142
QY 173 -OSKILYVGHSGQTINGLAAL-TMPEIVKMISSAALLCPISYLDHVS-----SFVLRAVAM 227
Db 143 GQEKIYVYGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPIATVKHAKSPCTKFL----- 197
QY 228 HLDQMLV--TMGIHOLNFRSDMGVQIVDSLCDGEHVD--CNNLLSAITGENC-CFNTSRI 282
Db 198 -LPDMWIKGLFGKKEFLYQTRFLRQLVYLQGVILDQICSNIMLLLGGFTNNMMSRA 256
QY 283 DYLYEVEPHPSSTKNLHLLFQMRKGTFAKYDYGL-LGNLRRYGHLPAPFDLSSIPESL 341
Db 257 SYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVRYRVDM--TV 314
QY 342 PIWMYGGDLADAVTDVORTIRELGSTPELLY---IGDYGHIDFVMSVKAKDDVYVDLI 398

Db 315 PTAMWTGGQDWSNPEDVKMLLSEV---TNLIYHKNIPEWAHVDFIWLGLDAPHRMYNEII 371

Qy 399 RFLRE 403

Db 372 HLMQ 376

RESULT 7

US-09-811-825-2

Sequence 2, Application US/09811825

Patent No. US20020144297A1

GENERAL INFORMATION:

APPLICANT: YAN, Chunhua et al.

TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL001170

CURRENT APPLICATION NUMBER: US/09/811.825

CURRENT FILING DATE: 2001-03-20

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 398

TYPE: PRT

ORGANISM: Human

US-09-811-825-2

Query Match 22.9%; Score 502; DB 10; Length 398;

Best Local Similarity 34.0%; Pred. No. 3.4e-40;

Matches 124; Conservative 67; Mismatches 146; Indels 28; Gaps 13;

Qy 55 GYPCTEHNVTKDGFLLSLQHIPHGKKAADSTG--PPVFLQHLGFGGDTWFINSAEQS 112

Db 32 GYPCEEYEVATEDGYILSVNRIPRGLVQ-PKKTGSRPVLLQHLGVLGGASNNISLNPNS 90

Qy 113 LGYILADNGFDWVGNVGRGTRWSKGHSFVSVDKLFWDWSHQELAEYDLLAMLGYYVTVT 172

Db 91 LGFILLADAGFDVWNGSRGNASRKHKTLSIDQDEFWAFSYDEMARDLPDAVINFILOKT 150

Qy 173 -QSKILYVGHSGQITMGLAAL-TMPEIVKMISSAALLCPISYLDHVA---SFLVRAYAM 227

Db 151 GQEKIYVGVYSGQTMGTAFSTMPDLAQKIMYFALAPIATVKHAKSPGTFKLL----- 205

Qy 228 HLDQMLV--TMGIHQLNFRSDMGVQIVDSLCDGEHVD--CNNLLSAITGENC-CFNTSRI 282

Db 206 -LPDMNKGFLGCKEFLYQTRFLRQLVLYLCGQVILQICSNIMLLGGFNTNNMMSRA 264

Qy 283 DYILEYEPHPSSTKNLHFLQMIKRGTFKDYDGL-LGNLRYGHLRPPAPDLSIPESL 341

Db 265 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNEKCNQPTPVRYVRDM--TV 322

Qy 342 PIWNGYGGDLADVDVQRTIRELGSTPELly---IGDYGHIDFVMSVKAKDDVYVDLI 398

Db 323 PTAMWTGGQDWSNPEDVKMLLSEV---TNLIYHKNIPEWAHVDFIWLGLDAPHRMYNEII 379

Qy 399 RFLRE 403

Db 380 HLMQ 384

RESULT 8

US-09-835-996A-21

Sequence 21, Application US/09835996A

Patent No. US20020142953A1

GENERAL INFORMATION:

APPLICANT: Ballinger, Dennis

APPLICANT: Loeb, Debra

APPLICANT: Montgomery, Julie

APPLICANT: Tang, Y. Tom

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Ryle

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhao, Qing

APPLICANT: Wehrman, Tom

APPLICANT: Dimanac, Radoje

APPLICANT: Ren, Feiyan

APPLICANT: Qian, Xiahong

APPLICANT: Wang, Dunrui

TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM

FILE REFERENCE: 28110/35915A

CURRENT APPLICATION NUMBER: US/09/835,996A

CURRENT FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: US 60/197,137

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: US 09/714,936

PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: US 09/667,298

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: US 09/631,451

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: US 09/598,042

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 45

SOFTWARE: PatentIn version 3.0

SEQ ID NO 21

LENGTH: 409

TYPE: PRT

ORGANISM: Homo sapiens

US-09-835-996A-21

Query Match 22.9%; Score 502; DB 10; Length 409;

Best Local Similarity 34.0%; Pred. No. 3.6e-40;

Matches 124; Conservative 67; Mismatches 146; Indels 28; Gaps 13;

Qy 55 GYPCTEHNVTKDGFLLSLQHIPHGKKAADSTG--PPVFLQHLGFGGDTWFINSAEQS 112

Db 43 GYPCEEYEVATEDGYILSVNRIPRGLVQ-PKKTGSRPVLLQHLGVLGGASNNISLNPNS 101

Qy 113 LGYILADNGFDWVGNVGRGTRWSKGHSFVSVDKLFWDWSHQELAEYDLLAMLGYYVTVT 172

Db 102 LGFILLADAGFDVWNGSRGNASRKHKTLSIDQDEFWAFSYDEMARDLPDAVINFILOKT 161

Qy 173 -QSKILYVGHSGQITMGLAAL-TMPEIVKMISSAALLCPISYLDHVA---SFLVRAYAM 227

Db 162 GQEKIYVGVYSGQTMGTAFSTMPDLAQKIMYFALAPIATVKHAKSPGTFKLL----- 216

Qy 228 HLDQMLV--TMGIHQLNFRSDMGVQIVDSLCDGEHVD--CNNLLSAITGENC-CFNTSRI 282

Db 217 -LPDMNKGFLGCKEFLYQTRFLRQLVLYLCGQVILQICSNIMLLGGFNTNNMMSRA 275

Qy 283 DYILEYEPHPSSTKNLHFLQMIKRGTFKDYDGL-LGNLRYGHLRPPAPDLSIPESL 341

Db 276 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNEKCNQPTPVRYVRDM--TV 333

Qy 342 PIWNGYGGDLADVDVQRTIRELGSTPELly---IGDYGHIDFVMSVKAKDDVYVDLI 398

Db 334 PTAMWTGGQDWSNPEDVKMLLSEV---TNLIYHKNIPEWAHVDFIWLGLDAPHRMYNEII 390

Qy 399 RFLRE 403

Db 391 HLMQ 395

RESULT 9

US-10-042-431-47

Sequence 47, Application US/10042431

Publication No. US20020182675A1

GENERAL INFORMATION:

APPLICANT: MCCARTHY, Sean A

APPLICANT: BARNES, Thomas M

APPLICANT: FRASER, Christopher C

APPLICANT: SHARP, John D

TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,

TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES

FILE REFERENCE: 10147-6U2

Db	83	LGFTLADAGFDVWNGSRGNWSRKHKHTLSIDQDEFWAFSYDEWAREDLPAVINFILOKT	14
Qy	173	-QSKILYVHGSHQGTIMGLAAL-TMPEIVKMISSAALLCPISYLDHVSA	218
Db	143	GOEKIYVVGYSQGTMGFIASFSTMPLELAQKIKMYFALAPATVVRHAKS	190

```
RESULT 12
US-10-056-744B-5
; Sequence 5, Application US/10056744B
; Publication No. US20020193303A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 58860, A HUMAN CHOLESTERYL ESTER
; FILE REFERENCE: HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MP2001-026P1RNM
; CURRENT APPLICATION NUMBER: US/10/056,744B
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/264,167
; PRIOR FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-056-744B-5

Query Match          12.08; Score 264; DB 9; Length 127;
Best Local Similarity 44.28; Pred. No. 3.9e-18;
Matches 53; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

Qy 104 WFINSAEQSLGYILLADNGFDVWIGNVRGTRWSKGHSTFSVHDKLFWDMWSQELAEYDLLA 163
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1  WIANGPNKSLAFILADACGYVWLGNSRGNTYSRKHKVKNLPSHSEFWDFSWHEMGMYDLPA 60

Qy 164 MLGYVYVVT-OSKILYVGHSGGTINGLAALT-MPEIVKMISAAALLCPISYLDHVSASFV 221
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 MIDYVLETTGOEKLHYVGHSGQTTFFVWMLSERPEYNEKIKSFHALAPVAYMKHVRSPV 120

RESULT 13
US-10-042-431-52
; Sequence 52, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; FILE REFERENCE: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; CURRENT APPLICATION NUMBER: 10147-6U2
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 52
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-52

Query Match          5.28; Score 114.5; DB 9; Length 144;
Best Local Similarity 24.68; Pred. No. 0.00089;
Matches 32; Conservative 30; Mismatches 59; Indels 9; Gaps 4;

Qy 278 NTSRIDYILEYEPHPSSTKNLHLFQMRKGTFAKYDGL-LGNLRRYGHLPAPAFDLSS 336
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6  NMSRSVYAAHTLAGTSVQNTLHWSQAVNSGELRFDWGSSETKLNKCNQPTPVRYVRD 65

Qy 337 IPESLPIMWGGGDLADVDVORTIRELGSTPELLY---IGDYGHIDFVMSYKAKDDV 393
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 M--TVPTAMWTGGQDLNSNPEDVKMLLSEV---TNLIYHKNIPEWAHVDFTWGLDAPHRM 120
```

```
Qy 394 YVDLIRFLRE 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 YNEIHLMOQ 130

RESULT 14
US-09-923-779-151
; Sequence 151, Application US/09923779
; Patent No. US20020076721A1
; GENERAL INFORMATION:
; APPLICANT: Pyte, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.553
; CURRENT APPLICATION NUMBER: US/09/923,779
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-779-151

Query Match          4.18; Score 90.5; DB 10; Length 465;
Best Local Similarity 20.48; Pred. No. 0.89;
Matches 80; Conservative 55; Mismatches 146; Indels 111; Gaps 22;

Qy 57 PCTPHNVETKDFLLSLQHIPHG-KNKAADSTG-----PPVFLOHGLF-QGGDTW 104
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 45 PWSPKDVNTR--FLLYTNENPNFQEAADSSISGSNFKTNRKTRFIHGFIDKGEENW 102

Qy 105 FINSAE-----QSLGYILLADNGFDVWIGNVRGTRWSKGHSTFSVHDKLFWDMWSQELAEY 159
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 103 LANVCKNLKFKVESVNCICVD-----WKGSR-----TGYTQASQNRIV---GAEVAYF 148

Qy 160 DLLAMLGYVYVVTQSKILYVGHs-----QCTINGLAALT-----MPEIV 198
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 149 --VEFLQSAFGYSFSPNVHIVGHSLGAHAAGAGRTNCTIGRTGLDPAEPCFGTPELV 206

Qy 199 KMISSAALLCPISYLDHV-----SASFVLRAVAMHLDQMLVTMGHQLNFRSDMGVQIVDS 254
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 207 RLDPDSAKFVDVIHTDGAPIVPNLGFSGMSQVVGHLDFPENGVMPCCKNILLSQIVD- 264

Qy 255 LCDG-----EHVDCNLLS-----AITGENC-----CF-----NTSR 281
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 265 -IDGIWEGTRDFAACNHLRSKYKYYTDSIVNPDGFAGFPFCASYNVFTANKCFPCSPGGCPQ 323

Qy 282 IDYYLEYEPHPSSTKNLHLFQMR--KGTFAKYDYG--LGNLRRYGHLPAPAFDLSS 336
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 324 MGYA--DRYEGKINDVGQKQFYLDGDSNFRARYKYSVYTLSCGKKVYTGCHILVSLFGNKG 381

Qy 337 IPESLPIMWG-----YGGDLADLADVDVQ 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 382 NSQVETPKGTLPDSTHSNEFSDVDVGDQ 413

RESULT 15
US-09-925-297-567
; Sequence 567, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
```

Tue Dec 31 09:21:28 2002

; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 567
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-567

Query Match 4.1%; Score 90.5; DB 10; Length 473;
Best Local Similarity 20.4%; Pred. No. 0.91;
Matches 80; Conservative 55; Mismatches 146; Indels 111; Gaps 22;

QY	57	PCTEHNVETKDGFLSLQHIPHG-KNKAADSTG-----PPVFLQHGFLF-QGGDTW	104
DB	53	PWSPKDVNTR--FLLYTNENPNFQEAADSSISGSNEKTNRKTRFIIHGFIIDRGEENW	110
QY	105	FINSAE-----OSLGYILADNGFDVWIGNVRCGRWSKGHSTFSVHDKLPWDHWSQELAEY	159
DB	111	LANVCKNLKFEVSNVICVD-----WKGGR-----TGYTQASONIRIV-----GAEVAYF	156
QY	160	DLMLGLVYVYVTQSKILYVGHG-----QGTIMGLAALT-----MPEIV	198
DB	157	--VEFLOSAGYSPSNVHVIGHSLGAHAAGEAGRRNTNGTIGRITGLDPAEPCFQGTPELV	214
QY	199	KMISSAALLCPISYLDHV-----SASFVLRAVAMHLDQMLVTMGIHQLNFRSDMGVQIVDS	254
DB	215	RLDPSDAKEVDVIHTDGAPIVNLGFGMSQVVGHLN-PPNGGVEMPGCKKNILSQIVD-	272
QY	255	LCOG-----EHVDCNNLLS-----AITGENC-----CF-----NTSR	281
DB	273	-IDGIWEGTRDFAACNHLRSYKYTDSIVNPDGFAGFPACASYNVFTANKCFPCPSGGCPQ	331
QY	282	IDYLEYEPHPSSTKNLHHLFOMIR--KGTFAKYDYL---LGNLRRYGHRLRPPAFDLSS	336
DB	332	MGHYA--DRYPGKTNDVGCKFYLDTDGASNFAWRWKYKSVTLGSKKVTGCHILVSLFGNKG	389
QY	337	IPESLPIMNG-----YGGLDLADVTDVQ	360
DB	390	NSKQYEIEFKGTLKPDSTHSNEFDSVDVGDLDQ	421

Search completed: December 30, 2002, 20:49:32
Job time : 71 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2002, 17:12:16 ; Search time 3834 Seconds
(without alignments)
11257.032 Million cell updates/sec

Title: US-09-699-652A-13
Perfect score: 1483
Sequence: 1 gcacgagtcacagcgcgcc.....ctctcaaaaaaaaaaaaaa 1483

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.un.*
- 28: em.vi.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	136.6	9.2	57564	8	AC006298	AC006298 Arabidops
2	100.8	6.8	793	8	AY048236	AY048236 Arabidops
3	94.2	6.4	2481	6	AX411234	AX411234 Sequence
4	94.2	6.4	2481	9	HSU08464	U08464 Human lysos
5	94.2	6.4	2481	9	HSU08464	U08464 Human lysos
6	94.2	6.4	2493	9	HSU08464	M74775 Human lysos
7	94.2	6.4	2586	9	BC012287	BC012287 Homo sapi
8	94.2	6.4	2626	9	HSLAL	X76488 H.sapiens m
9	92.6	6.2	2982	9	AK091558	AK091558 Homo sapi
10	89.4	6.0	1272	6	AX477670	AX477670 Sequence
11	88	5.9	1417	4	BOVME	L26319 Bos taurus
12	88	5.9	2577	10	BC031933	BC031933 Mus muscu
13	86.6	5.8	618	8	AY091711	AY091711 Arabidops
14	86.4	5.8	1228	6	AX392821	AX392821 Sequence
15	85.8	5.8	41805	3	CBRG44C02	AC084610 Caenorhab
16	83.8	5.7	417	6	AX150702	AX150702 Sequence
17	83.8	5.7	417	6	AX150702	AX150702 Sequence
18	80.4	5.4	2358	10	MMLYACLY	Z31689 M.musculus
19	80	5.4	1365	9	HSGLR	X05997 Human mRNA
20	80	5.4	1367	6	A01046	A01046 H.sapiens m
21	80	5.4	1367	6	A12714	A12714 Pregastric
22	78.6	5.3	41014	3	CEF01G10	Z81055 Caenorhabd
23	78.4	5.3	1336	6	A01157	A01157 R.norvegicu
24	78.4	5.3	1355	10	RNLIP	X02309 Rat mRNA fo
25	78.4	5.3	1378	6	A26689	A26689 Precursor o
26	78.4	5.3	1378	6	A26690	A26690 Precursor o
27	77.6	5.2	1211	3	DMLIPASE3	Y14367 Drosophila
28	77.6	5.2	25416	2	AC013100	AC013100 Drosophill
29	77.6	5.2	82393	3	AC003923	AC003923 Drosophill
30	77.6	5.2	194672	3	AC008360	AC008360 Drosophill
31	77.6	5.2	225668	3	AE003699	AE003699 Drosophill
32	72.8	4.9	2506	9	AK096406	AK096406 Homo sapi
33	71.8	4.8	16173	3	CEF54F3	Z79696 Caenorhabd
34	71	4.8	4543	2	AC017119	AC017119 Drosophill
35	71	4.8	96018	3	AE003783	AE003783 Drosophill
36	71	4.8	172748	3	AC018484	AC018484 Drosophill
37	71	4.8	186695	3	AC013351	AC013351 Drosophill
38	70.6	4.8	23525	2	AC116988	AC116988 Dictyoste
39	69.6	4.7	40114	3	AF067942	AF067942 Caenorhab
40	68.6	4.6	1473	6	AX392824	AX392824 Sequence
41	68	4.6	3144	10	S81497	S81497 lysosomal a
42	67.2	4.5	1185	6	AX472752	AX472752 Sequence
43	66.2	4.5	93932	3	AC004766	AC004766 Drosophill
44	66.2	4.5	175007	3	AC008342	AC008342 Drosophill
45	66.2	4.5	211784	2	AC017903	AC017903 Drosophill

ALIGNMENTS

RESULT 1
AC006298/c
LOCUS AC006298
DEFINITION Arabidopsis thaliana chromosome 2 clone F15A23 map mi398, complete sequence.
ACCESSION AC006298
VERSION AC006298.13
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 57564)
AUTHORS Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., VanAken,S.E.,

```

Barnstead,M.E., Mason,T.M., Bowman,C.L., Ronning,C.M.,
Benito,M.-I., Carrera,A.J., Creasy,T.H., Buell,C.R., Town,C.D.,
Nierman,W.C., Fraser,C.M. and Venter,J.C.
Unpublished
2 (bases 1 to 57564)
Lin,X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 57564)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
On Apr 18, 2002 this sequence version replaced gi:5598553.
FEATURES             location/Qualifiers
     source            1..57564
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /map="mi398"
                     /clone="F15A23"
     misc_feature      1..708
                     /note="overlap with BAC clone F27010
                     (AC007267:54979..55686)."
```

```

     gene              1458..4815
                     /gene="At2g15240"
                     /note="F15A23.2; supported by full length cDNA:
Ceres:120231"
                     join(1458..1623,2070..2137,2558..2583,2885..2975,
3078..3173,3541..3609,3810..3868,3974..4096,4191..4262,
4352..4815)
                     /gene="At2g15240"
     repeat_region     1477..1500
                     /rpt_family="GAA)n"
     CDS               join(1534..1623,2070..2137,2558..2583,2885..2975,
3078..3173,3541..3609,3810..3868,3974..4096,4191..4262,
4352..4416)
                     /gene="At2g15240"
                     /codon_start=1
                     /product="expressed protein"
                     /protein_id="AAD25568.2"
                     /db_xref="GI:20197743"
     gene              translation="MLPTTSRSRSSSSSRANPMFLQYFRIVKQOQMDVYTFWOML
NLTSPKVYQHTKYHKTQKNQARDPAFVICSLLVATVAVCVYDHSSHAVV
VYVSVLFTHLITGAVIATCCWFLTNSYLREETPNSHVVEQVWLYTFDVCSSFFP
MFVLVYVHYELSPLLIAHGFIPLLSNLFEMVGASYHYILNFLGYDVLPLELRTTF
LYPGVWIVLSPILISGFNPSRYFMNMYFSQRL"
                     complement(5001..7730)
                     /gene="At2g15230"
                     /note="F15A23.3"
     mRNA              complement(join(<5001..5514,5656..5722,6516..6659,
6739..6849,7152..7292,7389..7493,7587..>7730))
                     /gene="At2g15230"
     CDS               complement(join(5192..5514,5656..5722,6516..6659,
6739..6849,7152..7292,7389..7493,7587..7730))
                     /gene="At2g15230"
                     /codon_start=1
                     /product="putative lysosomal acid lipase"
                     /protein_id="AAD25569.1"
                     /db_xref="GI:4585908"
     gene              translation="MKWLLVAVLSTLFTSALTSHLLHSGPVNSLCADLIHPANYS
TEHSIQTKDGYILQRYVASIGPRLQGPVLLQGLPMAGDVNPLNSPKSLGFLA
DHGPDVWGNVGRYSGYCHVLTSDTKREFWDWQDLAMYDLAEMIOYLISNSKI
FLVGHQSQTALFRALQPHVAEMVEAAALLCIPISYLDHVTAPLVERVWFHLDQIEY
YLDVPEHPSSYKNIHLEFMIRKGTFAQYDYGYNKRLTGLSKPPEFILSHFASLP
MMWGGYDGTDLVDVHTLAEFLFSPPELLYLEDYGHIDFVLGSSAKEDYVKHMIOF
FRKVKSSW"
     repeat_region     complement(7968..8028)
     repeat_region     /rpt_family="AT_rich"
                     complement(9498..9533)
                     /rpt_family="AT_rich"

```

```

9883..9972
/rpt_family="AT_rich"
complement(10449..10521)
/rpt_family="AT_rich"
complement(10858..11720)
/gene="At2g15220"
/note="F15A23.4; predicted by genscan; supported by cDNA:
gi:13569549.gb.AF345341.1.AF345341"
complement(<10858..>11720)
/gene="At2g15220"
complement(11030..11707)
/gene="At2g15220"
/codon_start=1
/product="expressed protein"
/protein_id="AAD25570.1"
/db_xref="GI:4585909"
/translation="MTEHKIFFVLSMLVSLVNAVSVVDNSGDSGTRFRREIG
GISYQTQLRSATDEVRWLRFTQNPDRKSVTKITLFMENGDGVAYNANETHEFNVG
LAGYSGDVKREFTGVVYHEVHSMQWNGAGRAPGLIEGIADYVRLKAGYAPSHWGP
GRGDRWDGQYDVTARFLDYCNGLRNGFVAELNKKMRNGYSQGFVVDLLGKDVNQLWRE
YKAKYQ"
12211..12236
/rpt_family="AT_rich"
12536..47776
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomously DNA-transposon
ARNOLD3."
12536..28026
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomously DNA-transposon
ARNOLD4."
12550..28026
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomously DNA-transposon
ARNOLD1."
12585..13525
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomously DNA-transposon
ARNOLD3."
12906..13527
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomously DNA-transposon
ARNOLD4."
13151..13431
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomously DNA-transposon
ARNOLD1."
13157..54188
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomously DNA-transposon
ARNOLD3."
13157..13441
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomously DNA-transposon
ARNOLD3."
13157..13357
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomously DNA-transposon
ARNOLD3."
13157..13347
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomously DNA-transposon
ARNOLD1."
13157..13263
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomously DNA-transposon
ARNOLD1."
13159..13527
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomously DNA-transposon
ARNOLD4."
13159..13443
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomously DNA-transposon
ARNOLD4."
13159..13359
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomously DNA-transposon
ARNOLD4."
13241..13496
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomously DNA-transposon
ARNOLD1."
13243..13527
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomously DNA-transposon
ARNOLD4."
13243..13525
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomously DNA-transposon

```



```
repeat_region      ARNOLD3."
13326..13525
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomous DNA-transposon
ARNOLD1."
13327..13527
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomous DNA-transposon
ARNOLD4."
13327..13525
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomous DNA-transposon
ARNOLD3."
13410..13525
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomous DNA-transposon
ARNOLD1."
13411..13525
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomous DNA-transposon
ARNOLD3."
13966..14153
/rpt_family="Rf:ARNOLD2|ARNOLD2 Autonomous DNA-transposon
ARNOLD2."
14080..14197
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomous DNA-transposon
ARNOLD4."
14080..14153
/rpt_family="Rf:ARNOLD3|ARNOLD3 Autonomous DNA-transposon
ARNOLD3."
14090..14197
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomous DNA-transposon
ARNOLD4."
14096..14153
/rpt_family="Rf:ARNOLD4|ARNOLD4 Autonomous DNA-transposon
ARNOLD4."
14122..14197
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomous DNA-transposon
ARNOLD1."
14124..14197
/rpt_family="Rf:ARNOLD2|ARNOLD2 Autonomous DNA-transposon
ARNOLD2."
14124..14190
/rpt_family="Rf:ARNOLD1|ARNOLD1 Autonomous DNA-transposon
ARNOLD1."
14140..14290
repeat_region
```

Query Match 9.2%; Score 136.6; DB 8; Length 57564;

Best Local Similarity 63.0%; Pred. No. 2.le-19;

Matches 211; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 933 TTTTTCAGATGATCAGGAAGGCACTTTCGCAAGTATGACTATGGTTATTGGGAACC 992

Db 5523 TTACTCCAGTGATTCGGAAGGGAACCTTTGCAATACCACTACCGGTATTTCAAAAATC 5464

QY 993 TAAGCGCTACGGTCTTCGTCCTCCGCAATTTGACCTAAGCAGCATACCAGAAATCAC 1052

Db 5463 TACGGACTTATGGCTGTGGAACCTCCAGAAATTCATACCTAAGCCACATCCCGGCATCAT 5404

QY 1053 TGCCCATATGGATGGATATGAGGTCTTGATGCTATGGCTGATTAACCGATGTTCCAG 1112

Db 5403 TACCAATGTGGATGGGTATGTTGGAACCTAGTGGTTTAGCAGATGTGACAGATGTGGAAC 5344

QY 1113 GTACTATCAGAGCTGGGATCTACACCAAGTCTCTGTACATTTGCTGACTATGCCCAT 1172

Db 5343 ATACTCTCCGGAACACTACCTCCAGTCCAGATTTACTATATCTTGAGGATTTGTCACA 5284

QY 1173 TTGATTTTGTATGACGCTGAAGCGGAAGATGATGTTTATGTGGACCTTAATAAGATTTC 1232

Db 5283 TTGACTTTTGTGCTTGGCTCAAGTCTAAGTAAGATGCTCTATAAGCACATGATTCATT 5224

QY 1233 TTAGGGAATGGATGGCATATAGCTATTAGGAT 1267

Db 5223 TCAGAGCAAGGTTAAGTCTAGTAGTTAGTTGTTGTAATAAT 5189

RESULT 2
A1048236

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY048236
Arabidopsis thaliana AT5g14180/MUA22_18 mRNA linear PLN 05-AUG-2001
AY048236
AY048236.1 GI:15081687
FLI_CDNA.
Arabidopsis thaliana.
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 793)
Koesema,E., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Shinn,P.,
Banh,J., Bowser,L., Carninci,P., Dale,J.M., Gibson,H.A.,
Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T.,
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J.,
Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S.,
Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Yu,S.,
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 793)
Koesema,E., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Shinn,P.,
Banh,J., Bowser,L., Carninci,P., Dale,J.M., Gibson,H.A.,
Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T.,
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J.,
Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S.,
Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Yu,S.,
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission

TITLE
JOURNAL

Submitted (09-JUL-2001) Salk Institute Genomic Analysis Laboratory
(SIGnAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Koesema,E., Chen,H.,
Cheuk,R., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Koesema,E. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source

Location/Qualifiers
1..793
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAFL05-14-M10(R13071)"
/note="ecotype: Columbia"
72..689
/note="unknown protein"
/codon_start=1
/product="AT5g14180/MUA22_18"
/protein_id="AAK82499.1"
/db_xref="GI:15081688"
/translation="MLSPVAVLSHMTTIVGDIAAKTFLAEATISLWPEFNPKSLGV
DFIKAIICLAGIDCYDLVITGKNCLNASTIDFLANEPQSTSTKNIHLAOTVRD
KELRYNYGSSDRNIKKHVGQAIIPPAYNISAIPELPLFFSYGGLDLSADVKDVEFLLD
QPKYHDIDKMNQVFKDFAHADFIMGVTAQDVVYNQVATFFRQA"

5'UTR
CDS

3'UTR
BASE COUNT

228 a 170 c 166 g 229 t

Query Match 6.4%; Score 94.2; DB 9; Length 2626;
Best Local Similarity 55.2%; Pred. No. 2.9e-10;
Matches 249; Conservative 0; Mismatches 193; Indels 9; Gaps 3;

QY 194 GGCTACCCCTGCACGAGCAGACAGTGAACAAAGATGGATTCCTTTTATCTCTTCAG 253
DB 278 GGATTCCTAGTGAAGAAATACCTAGTTGAGACAGAAGATGGATATATCTGTGCCTTAAC 337
QY 254 CATATCCCATGCGCAAAATAAAGCAGACAGATAGTACTGGCCCTCCAGTT---TTTCTT 310
DB 338 CGAATTCCTCATGGGAGGAAGAACCAATCTGACAAAGTCCCAAAACAGTTCTCTCTCTG 397
QY 311 CAACATGGTCTTTTTCAGGAGGAGACACATGGTTCAATAAATCTGCTCAGCAATCACTT 370
DB 398 CAACATGGTCTGCTGCGCATCTAGTACTGGTTCACAAACCTTGCCACAGCAGCCTG 457
QY 371 GGTATATCTCTGATACAGCGTTTGTGATTTGGATTGGAAATGTCGGTGGACAGCGT 430
DB 458 GGTCTCATCTCTGATGCTGCTGTTTACGTTGTGGATGGCAACAGCAGAGAAATACC 517
QY 431 TGGAGTAAAGGTCATCAACCTTTTCTGTTTCATGATAAGCTTTTCTGGGATTTGGAGCTGG 490
DB 518 TGTCTCGGAACATAAGACATCTCAGTTTCTCAGGATGAATCTGGGCTTTCAGTTAT 577
QY 491 CAAGAGTTAGCTGAATATGACCTTTTAG---CAATGCTAGGCTATGTTATACAGTCACA 547
DB 578 GATGAGATGGCAAAATATGACCTACAGCTTCCATTAATCTTCTGAATAAATACG 637
QY 548 CAGTCCAAATCTATATCTGGGGCATTCACAGGAACTATATGAGTTGGCGTTGGCGTTTG 607
DB 638 CAAGAACAAAGTATATATGTTGGTCTATCTCAAGGCACCACCTATAGTTTATAGCATTT 697
QY 608 AC---GATCCCGGAAATAGTAAATGATTA 635
DB 698 TCACAGATCCCTGAGCTGGCTAAAGGATTA 728

RESULT 9
AK091558
LOCUS Homo sapiens cDNA FLJ34239 fis, clone FCBBF3027755, highly similar
to LYSOXOMAL ACID LIPASE/CHOLESTERYL ESTER HYDROLASE PRECURSOR (EC
3.1.1.13).
ACCESSION AK091558
VERSION AK091558.1 GI:21749961
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens fetal brain cDNA to mRNA, clone_lib:FCBBF3
clone:FCBBF3027755.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chiba,Y., Suigiya,T., Irie,R., Otsuki,T., Sato,H.,
Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsuo,K.,
Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2982)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

source
1..2982
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FCBBF3027755"
/tissue_type="brain"
/clone_lib="FCBBF3"
/dev_stage="fetal"
/note="cloning vector: pME18SFL3"

BASE COUNT 847 a 596 c 615 g 924 t
ORIGIN

Query Match 6.2%; Score 92.6; DB 9; Length 2982;
Best Local Similarity 55.0%; Pred. No. 6.8e-10;
Matches 248; Conservative 0; Mismatches 194; Indels 9; Gaps 3;

QY 194 GGCTACCCCTGCACGAGCAGACACCTTGAACAAAGATGGATTCCTTTTATCTCTTCAG 253
DB 810 GGATTCCTAGTGAAGAAATACCTAGTTGAGACAGAAGATGGATATATCTGTGCCTTAAC 869
QY 254 CATATCCCATGTCGCAAAATAAAGCAGCAGATAGTACTGGCCCTCCAGTT---TTTCTT 310
DB 870 CGAATTCCTCATGGGAGGAAGAACCAATCTGACAAGGTCCCAAAACAGTTGCTCTTCCTG 929
QY 311 CAACATGTCCTTTTTCAGGAGGAGACACATGGTTTCATAAACTCTGCTGAGCAATCACTT 370
DB 930 CAACATGGCTTCTGCTGGCAGATTCTAGTAACCTGGTGCACAAACCTTGCCACGAGCAGCCTG 989
QY 371 GGTATATCTCTGTCGATAACGGTTTTCATTTGATTTGGATTGGAAATGTCGGTGGACGGCT 430
DB 990 GGTCTCATCTCTGTCGATGCTGTTTTCACGCTGGATGGGCAACAGCAGAGAAATACC 1049
QY 431 TGGAGTAAAGGTCATCAACCTTTTCTGTTTCATGATAAGCTTTCTCTGGGATTTGGAGCTGG 490
DB 1050 TGGTCTCGAAACATCAGACACTCTCAGTTCTCAGGATGAATCTGGGCTTTCAGTTAT 1109
QY 491 CAAGAGTTAGCTGAATATGACCTTTTAG---CAATGCTAGGCTATGTTATACAGTCACA 547
DB 1110 GATGAGATGGCAAAATATGACCTACACCTCCATTAATCTCATTGAATAAATACG 1169
QY 548 CAGTCCAAATCTATATGTTGGGCAATTCACAGGAACTATATGAGTTGGCGGCTTTG 607
DB 1170 CAAGAACAAAGTATATGTTGGTCAATTCCTCAAGGCACCACCTATAGTTTATAGCATTT 1229
QY 608 AC---GATGCCCGAAATAGTAAATGATTA 635
DB 1230 TCACAGATCCCTGAGCTGGCTAAAGGATTA 1260

RESULT 10
AX477670
LOCUS AX477670 1272 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 8 from Patent WO0246418.
ACCESSION AX477670
VERSION AX477670.1 GI:22216841
KEYWORDS human.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Griffin,J.A., Gandhi,A.R., Ramkumar,J., Tang,Y.T., Ding,L., Yue,H.,
Gietzen,K.J., Sapperstein,S.K., Honchell,C.D., Bruns,C.M.,
Duggan,B.M., Xu,Y. and Lee,S.

TITLE

Lipid-associated molecules

JOURNAL

Patent: WO 0246418-A 8 13-JUN-2002;

FEATURES

Location/Qualifiers

```
source
1..1272
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7483978CB1"
BASE COUNT      351 a      263 c      311 g      347 t
ORIGIN

Query Match      6.0%; Score 89.4; DB 6; Length 1272;
Best Local Similarity 53.6%; Pred. No. 3e-09;
Matches 233; Conservative 0; Mismatches 196; Indels 6; Gaps 2;

QY 183 TCCTGCCACTCGGCTACCCCTGCACCGACACACAGCTTGAACAAAGATGATTCCTTT 242
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 TCAACATCAAGCTATCCCTCTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCC 196
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 243 TATCTCTTCAGCATATCCACATGCGCAAAATAAAGCAGCAGATA---GTACTGGCCCTC 299
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 TTTCTGTTAACAGGATTCCTCGAGGCTAGTGCACCTAAGACAGAGTTTCAGGCCCTG 256
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 300 CAGTTTCTTCAACATGTCCTTTTCAGGGAGGAGACACATGGTTTCATAAACTCTGCTG 359
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 TGGTGTACTGAGCATGCGCTAGTTGGAGGTGCTAGCAACTGGATTTCCCAACCTGCCCA 316
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 360 AGCAATCACTGGGTATATCCTTCGTGATAACGGTTTTCATGTTGGATGGGAATGTCC 419
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 ACAATAGCTGGGCTTCATCTGGCAGATGCTGGTTTTCACGTGTGGATGGGGAACAGCA 376
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 420 GTGGAAAGCTTGGAGTAAAGTTCATCAACCTTTTCTGTTTCATGATAAGCTTTTCTGGG 479
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 GGGAAACCCCTGGTCTCGAAACACAGACACTCTCCATAGACCAGATGAGTCTGGG 436
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 480 ATTGGAGCTGGCAAGAGTAGTGAATATGACCTTTTAGCAATGCTAGGCT---ATGTGT 536
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 437 CTTTCAGTTATGATGAGATGGCTAGTGTGACCTTCTCCAGTGATAAACTTATTTTGC 496
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 537 ATACACTCACAGTCCAAATCTATATGTGGGGCATTCACAGGGAACATAAATGGGTT 596
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 497 AGAAACGGCCGAGGAAAGATCTATTATGTGGCTATTTCACAGGCGACCAACCATGGCT 556
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 597 TGGCGGCTTTGACGA 611
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 557 TTATTGCATTTTCCA 571
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
BOVME
LOCUS      BOVME
DEFINITION Bos taurus pregastric esterase mRNA, complete cds.
ACCESSION L26319
VERSION    L26319.1 GI:600756
KEYWORDS   esterase; pregastric esterase.
SOURCE     Bos taurus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
           Bovinae; Bos.
REFERENCE 1 (bases 1 to 1417)
AUTHORS    Timmermans, M.Y., Teuchy, H. and Kupers, L.P.
TITLE      The cDNA sequence encoding bovine pregastric esterase
JOURNAL    Gene 147 (2), 259-262 (1994)
MEDLINE    95011625
PUBMED     7926811
FEATURES   Location/Qualifiers
           1..1417
           /organism="Bos taurus"
           /db_xref="taxon:9913"
           /cell_type="Secretory cell"
           /tissue_type="Tongue serous gland"
           /dev_stage="25 weeks"
           73..1266
           /EC_number="3.1.1.3"
           /codon_start=1
           /product="pregastric esterase"

CDS
1..1272
/protein_id="AAA57037.1"
/db_xref="GI:600757"
/translation="MWLLVTVCFTHMSGNAFCFLGKIAKNPEAMNVSQMSLTSYWGYP
SEMKVITADCVILQVYRIPHGKNNHNGORPVVFLQGLLGSATNWSLNPKNSLG
ELLADAGDYVWLGNSRGNTWAQEHLYSPDSFEWAFSEDEMAEYDLESTIDFLIRRT
GOKLHYVGHSGGTTIGTAFSTLAERIKVFYALAPVATVATVYKTKSLFNKLALIPH
FLFKIIFGDKMFYPHTLEQFLGVEMCSRETLDVLCKNALFAITGVNDKNMNSRLDY
YIAHNPAGTSVQNTLHWRQAVKSGKQAFDMGAPYQNLHMYHPTPPYLYNTAMNVP
I AWSADNDLLADPDQDFLLSKLSNLVYKKEIPNYNHLDFIWMADPAQEVYNEIVSLM
AEDKK"
sig_peptide      73..129
mat_peptide      130..1263
/product="pregastric esterase"
/EC_number="3.1.1.3"
/evidence="triacylglycerol acylhydrolase"
/evidence="experimental"
1378..1385
polyA_signal
polyA_site       1396
BASE COUNT      408 a      306 c      276 g      427 t
ORIGIN

Query Match      5.9%; Score 88; DB 4; Length 1417;
Best Local Similarity 53.8%; Pred. No. 6.2e-09;
Matches 228; Conservative 0; Mismatches 190; Indels 6; Gaps 2;

QY 194 GGCTACCCCTGCACCGACACAAAGTTGAACAAAGATGGATTCCTTTTATCTCTCTCAG 253
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 196 GGCTACCCAAAGTGAGATGCATTAAGTTATACTGGGATGGTTATATCTCTCAGGCTCAT 255
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 254 CATATCCCAACATGGCAAAATAAAGCAGCAGAT---AGTACTGGCCCTCCAGTTTTCCT 310
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 256 CGGATTCTCTATGAAAGAAATAATGCTAATCATTTAGGTCAGAGACCTCTGTTGTTCTG 315
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 311 CAACATGGTCTTTTTCAGGGAGGAGACACATGGTTTCATAACTCTGCTCAGCAATCATCT 370
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 316 CAGCATGGTCTTCTTGGATCAGCCACAAACTGGATTTCCAACTGCCCCAAGAACAGCCTG 375
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 371 GGGTATATCTTCTGCTGATAACCGTTTGTGATTTGGATTGGGAATGTCGCTGGAACGCGT 430
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 376 GGCTTCTCTGCGAGATGCTGTTATGACGTGTGCTGGGACAGACAGAGAAACACC 435
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 431 TGGAGTAAAGGTCAITTCACCTTTTCTGTTTCATGATAAGCTTTTCGGAATGGAGCTGG 490
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 436 TGGGCCCAAGAACTTTTATCTATTCACCAGACTCCCCGGAATTTCTGGGCTTTACGCTTT 495
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 491 CAAGAGTTAGCTGAATACACCTTTTAGCAATGCTAGGCTATGTATACAGTCAC--A 547
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 496 GATGAATGCGGAATATGACCTTCCATCTACAATTGATTTCATCTTAAGGAGAACAGGA 555
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 548 CAGTCCAAAATTTCTATATGTGGGCGATTTCACAGGGAACATAATGGGTTTGGCGGCTTTG 607
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 556 CAGAAGAAGCTACACTATGTGGCCATTCCCAAGGACCAACCATTTGGTTTATCGCCTTT 615
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 608 ACGA 611
    |||
Db 616 TCTA 619
    |||

RESULT 12
BC031933
LOCUS      BC031933
DEFINITION Mus musculus, RIKEN cDNA 4632427C23 gene, clone MGC:25790
IMAGE:4020396, mRNA, complete cds.
ACCESSION BC031933
VERSION    BC031933.1 GI:21594465
KEYWORDS   MGC.
SOURCE     house mouse.
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 2577)
AUTHORS    Strausberg, R.
TITLE      Direct Submission
```


FEATURES source Location/Qualifiers
1..618
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U13071"
/note="This clone is in pUNI 51
ecotype: Columbia"
1..618
/note="unknown protein"
/codon_start=1
/product="AT5g14180/MUA22_18"
/protein_id="AA010310.1"
/db_xref="GI:20147187"
/translation="MLSPVAYLSHMTTVIGDIAAKTFLAEATSTILGWPEFNPKSLVIG
DFKAIICLKIDCVSVITGKNCCLNASTIDLFLANPQSTSTKNMHLAQTVRD
KELKYNKSGSDRNKIKHYGAIRPAYNISAIHELPLFFSYGLGSLADKVDVEFLID
QFYHIDIKMNQFVKDIAHAFIMGTAKDVVYNQVATEFKQA"
BASE COUNT 175 a 144 c 130 g 169 t
ORIGIN
Query Match 5.8%; Score 86.6; DB 8; Length 618;
Best Local Similarity 49.78; Pred. No. 1.le-08;
Matches 286; Conservative 0; Mismatches 274; Indels 15; Gaps 2;
Qy 654 GTCCATTCTTATCTTGATCATGTTAGTGTAGTGTGTTTCTTCAGAGCAGTCGCCATGC 713
Db 8 GTCCCGTTGCTTATCTCAGCCACATGACCACCGTCATCGCGGATATCCGCCCAAAACCT 67
Qy 714 ATCTTGATCAGATCGTGTACTATGGAATTCACCAGCTGAACCTCCGTAGCGACATGG 773
Db 68 TCCTCGCGCGGCCACCTCTATCTTGTGATGCCAGATTTAAACCGAAAAGTGATAG 127
Qy 774 GGGTTCAAAATAGTAGATCTTTGTGCGATGTGAACACGTGGATTGCAACAATTTGCTAT 833
Db 128 TAGGGATTTTATAAGCTATATGCCTTAAAGCGGGGATCGACTGTTATGATTAGTCT 187
Qy 834 CTGGGATTACAGGGGAAAACCTGTTCTTCAATACATCAAGGATTGATTATTATTGGAGT 893
Db 188 CTGTTATCACTGGAAAGAAATTCGTGCTTAAACGCATCAACCATTTGCTCTCTTGCCA 247
Qy 894 ATGAACCTCATCCATCATCGACAAAAAATCTGCACCATCTTTTCAGATGATCAGGAAG 953
Db 248 ACGAACCTCAATCTACTTCCACCAAGAACATGATCCACCTTGCTCAACCGTAAGACACA 307
Qy 954 GCACCTTCGCAAGATGATCACTATGGTTATTGG---GAAACCTAAGCGCTACGGTCATT 1010
Db 308 AGGAGTTGAGAAAATACAACTATGGAACAGTGTATCGCAACATAAACATTACGGTCAAG 367
Qy 1011 TCGGTCTCCCGCATTTGACCTAAGCAGCATACCAGAACTACTGCCCATATATGGATGGAT 1070
Db 368 CGATACCAACCGCATACAACATATCGGGGATCCACACGAGCTCCCGCTTTTTCAGCT 427
Qy 1071 ATGAGGTCTTGATGCAATGCGTGNATGTAACCGATGTTACCGGTACTATCAGAGCTGG 1130
Db 428 ACGGTGTTTATAGTAGCGCTGCTGATGTAAGAGCGTTGAGTTCTCTCTTGACCACTTA 487
Qy 1131 GAT-----CTACACCAGAACTTCTGTACATTTGTTGCTATGCGCATATTTGATT 1178
Db 488 AGTATCAGCAGATCGATGAAGATGATGTGCAGTTTGTGAAGACTATGCTCATGCTGATT 547
Qy 1179 TTGTTATGAGCGTGAAGCGGAAAGATGATGTTTAT 1213
Db 548 TCATCATGGGTGTTACTGCCCAAGACGTTGTTTAT 582
RESULT 14
AX392821
LOCUS
DEFINITION Sequence 7 from Patent W00216597.
ACCESSION AX392821
VERSION AX392821.1 GI:19700917
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Griffin,J.A., Patterson,C., Gandhi,A.R., Lu,Y., Yao,M.G.,
Baughn,M.R., Walla,N.K., Hafalia,A.J., Ding,L., Tribouley,C.M.,
Das,D., Thornton,M. and Lal,P.
TITLE Lipid metabolism enzymes
JOURNAL Patent: WO 0216597-A 7 28-FEB-2002;
FEATURES Location/Qualifiers
source 1..1228
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7479063CB1"
BASE COUNT 376 a 244 c 259 g 349 t
ORIGIN
Query Match 5.8%; Score 86.4; DB 6; Length 1228;
Best Local Similarity 53.58; Pred. No. 1.4e-08;
Matches 227; Conservative 0; Mismatches 191; Indels 6; Gaps 2;
Qy 194 GGTACCCCTTGCACGAGCACAACGTTGAAACAAAGATGATTCCTTTTATCTCTTCAG 253
Db 145 GGCTACCCCGAGTGAAGAGTATGAAGTCAACCTGAGATGGGTATATCTCTTGTCAAC 204
Qy 254 CATATCCACATGCAAAATAAAGCAGCAGATAGTACTGGCCCTCCAGTT---TTTTCTT 310
Db 205 AGAATTCCTTATGGCGAACACATGCTAGGAGCAGAGTCCCGGCCAGTTGTATATG 264
Qy 311 CAACATGTCTTTTTCAGGGAGGAGACACATGTTTCATAAACTCTGCTGAGCAATCACTT 370
Db 265 CAGCATGCCCTGTTTGCAGGACAGCCTTACTGGCTTGAGAAATATGCCAATGGAAGCCTT 324
Qy 371 GGGTATATCTTGTGTGATACGGTTTGTGATTTGGATTGGAAATGTCGGTGGAAACGGT 430
Db 325 GGATTCTCTTACGAGATGCGAGTTATGATGATGGTGGAAACAGTCGGGGAACACT 384
Qy 431 TGGAGTAAAGTCTATCAACCTTTTCTCTTCATGATAGCTTTTCTGGGATTTGGAGTGG 490
Db 395 TGGTCAAGAACACAAACACTCTCAGAGACAGATGAGAAATCTGGGCCCTTTAGTTTT 444
Qy 491 CAAGAGTTAGCTGAATATGACCTTTTAGCAATGCTAGGCT---ATGCTGATACAGTCACA 547
Db 445 GATGAATGGCCAAATATGATCTCCAGAGTATAGACTTCATTGTAATAAAACTGGT 504
Qy 548 CAGTCCAAAATTTCTATATGTTGGGCAATTCACAGGGAACATAATAGGGTTTGGCGGCTTTG 607
Db 505 CAGGAGAAATTTGATTTTCATTGGACATTCCTTGGCACTACAAATAGGGTTTGTAGCCCTT 564
Qy 608 ACGA 611
Db 565 TCCA 568
RESULT 15
CBRG44C02/C
LOCUS CBRG44C02 41805 bp DNA linear INV 04-NOV-2000
DEFINITION Caenorhabditis briggsae cosmid G44C02, complete sequence.
ACCESSION AC084610
VERSION AC084610.1 GI:11095059
KEYWORDS HTG.
SOURCE
ORGANISM Caenorhabditis briggsae.
Caenorhabditis briggsae
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 41805)
AUTHORS Washington University Genome Sequencing Center.
TITLE The C. briggsae Genome Sequencing Project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 41805)
AUTHORS Wilson,R.

TITLE The sequence of C. briggsae cosmid G44C02
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 41805)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (04-Nov-2000) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA
e-mail: jsplie@watson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

FEATURES
source
1. .41805
Location/Qualifiers
/organism="Caenorhabditis briggsae"
/strain="Gujarat G16"
/db_xref="taxon:6238"
/clone="G44C02"

BASE COUNT 13098 a 8049 c 8106 g 12552 t
ORIGIN

Query Match 5.8%; Score 85.8; DB 3; Length 41805;
Best Local Similarity 54.7%; Pred. No. 3.6e-08;
Matches 193; Conservative 0; Mismatches 157; Indels 3; Gaps 1;
Qy 242 TTAATCTCTTACAGCATATCCACATGGCAAAATAAAGCAGCAGATAGTACTGGCCCTCCA 301
Db 29801 TTGTCCTATGCCCTAATACCATATTCAGAGAAATGCCATACAAAGTCAAAACGTCCAATT 29742
Qy 302 GTTTTCTTCAACATGGTCTTTTTCAGGGAGGACACATGGTTTCATAAAACTCTGCTGAG 361
Db 29741 GTATTTTACAAATGGTCTCTAGCAGATGGATTTTCATGGATTCCAAATTTGGCTAAT 29682
Qy 362 CAATCACTTGGGTATATCCTTGCTGTATACCGTTTTCATGTTGGATGGGAATGTCGGT 421
Db 29681 CAATCAGCTGGATTTGTTTTTGGCGATGCGAGGGTTTGATGTGGATTGCCAAATAGTCGG 29622
Qy 422 GGAACGGCTTGGAGTAAAGGTCATTCAACCTTTTCTGTTCATGATAAGCTTTTCTGGGAT 481
Db 29621 GGTACTCCAGCATCACAAAGCATATCGGATATGCTCCAGAAATCAGAAATCTGGAAT 29562
Qy 482 TGGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTAGCAATGCTAGGCTATGCTATACA 541
Db 29561 TTCATATGGCAACAAATGCTCTGAATAGCATTTGACTTCGAGTGTGATTACGCTCTGAAA 29502
Qy 542 GTCACACAGTCCAAATCTA--TATGTGGGCAATTCACAGGGAACATAAT 591
Db 29501 GAGACAAACAAAGTTTTTATACTATTTGGGACATTTCTCAAGGAACATATGAT 29449

Search completed: December 30, 2002, 18:29:11
Job time : 3950 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2002, 17:10:06 ; Search time 281 Seconds
(without alignments)
11885.101 Million cell updates/sec

Title: US-09-699-652A-13
Perfect score: 1483
Sequence: 1 gcacgagtacacgagcgcgc.....ctctcaaaaaaaaaaaaaa 1483

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1483	100.0	1483	21	AAZ34956 Rice acid triacylg
2	529.4	35.7	859	21	AAZ34953 Corn acid triacylg
3	445.6	30.0	1438	21	AAZ34959 Soybean acid triac
4	184.6	12.4	1718	21	AAZ34958 Soybean acid triac
5	94.2	6.4	802	22	AAI97406 Human neuroblastom
6	94.2	6.4	2481	22	ABN97383 Gene #3881 used to
7	94.2	6.4	2493	22	AAH02906 Human shear stress
8	91.2	6.1	1411	24	ABL57144 Kid goat pregastric
9	91.2	6.1	8324	24	ABL57145 Yeast YE-1 express

10	89.4	6.0	1116	24	ABK86570 Human cDNA encodin
11	89.4	6.0	1210	24	ABK86569 Human cDNA encodin
12	89.4	6.0	1269	22	AAF45132 Human TANGO 294 OR
13	89.4	6.0	1384	22	AAI19220 Human CGI162 (or C5
14	89.4	6.0	2035	22	AAI19226 Human CGI162 (or C5
15	89.4	6.0	2044	22	AAF45131 Human TANGO 294 CD
16	86.4	5.8	1228	24	ABK12385 cDNA encoding huma
17	83.8	5.7	417	22	AAH50931 Lipid degradation
18	83.8	5.7	417	22	AAH50931 P patens lipid met
19	83.2	5.6	1284	24	ABK85978 DNA encoding human
20	81.6	5.5	1360	24	AAI47480 Human lysosomal ac
21	81.6	5.5	1695	22	AAH57420 Human stomach cell
22	80	5.4	617	21	AAA44349 Human secreted exp
23	80	5.4	1367	7	AAH60685 Sequence encoding
24	80	5.4	1367	7	AAH60566 Human gastric lipa
25	80	5.4	1367	17	AAH58916 Catalpa acid triac
26	79.6	5.4	705	21	AAZ34952 RGL precursor. Or
27	78.4	5.3	1378	14	AAQ42310 Drosophila melanog
28	77.6	5.2	1211	23	ABL29771 Drosophila melanog
29	77.6	5.2	3442	23	ABL29770 Rat prelingual lip
30	76.8	5.2	1336	6	AAH50385 Drosophila melanog
31	71	4.8	1242	23	ABL06473 Drosophila melanog
32	71	4.8	4196	23	ABL05472 Drosophila melanog
33	68.6	4.6	1473	24	ABK12388 cDNA encoding huma
34	67.2	4.5	1185	24	AAH31193 Human triacylglyce
35	66.2	4.5	1197	23	ABL12389 Drosophila melanog
36	66.2	4.5	3269	23	ABL12388 Drosophila melanog
37	65.8	4.4	1098	24	AAH27802 Human lysosomal ac
38	65.4	4.4	1200	22	AAH78205 Nucleotide sequenc
39	65.2	4.4	735	22	AAI24482 Human breast canc
40	64.4	4.3	1528	15	AAQ68388 Canine gastric lip
41	64.4	4.3	1528	17	AAH58915 Dog gastric lipase
42	64.4	4.3	1531	15	AAQ68389 Canine gastric lip
43	62.4	4.2	1194	24	AAH27803 Human lysosomal ac
44	60.8	4.1	699	22	AAH28679 Human protein hp03
45	60.8	4.1	828	24	AAH31194 Human triacylglyce

ALIGNMENTS

RESULT 1
AAZ34956
ID AAZ34956 standard; cDNA; 1483 BP.
XX
AC AAZ34956;
XX
DT 28-FEB-2000 (first entry)
XX
DE Rice acid triacylglycerol lipase cDNA.
XX
KW Triacylglycerol lipase; rice; fatty acid; seed oil;
KW vegetable oil; transgenic plant; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 32..1264
FT /*tag= a
XX
PN WO9955883-A2.
XX
PD 04-NOV-1999.
XX
PF 29-APR-1999; 99WO-US09280.
XX
PR 30-APR-1998; 98US-0083688.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
XX WPI; 2000-062036/05.
DR

PN WO9955883-A2.
 XX 04-NOV-1999.
 XX 29-APR-1999; 99WO-US09280.
 XX 30-APR-1998; 98US-0083688.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
 PI WPI: 2000-062036/05.
 DR P-PSDB; AAY32309.
 XX
 XX Novel plant triacylglycerol lipase polynucleotides used to alter the
 PT level of the enzyme in transgenic plants
 XX
 XX Claim 2: Page 48; 65pp; English.
 XX
 CC This is the nucleotide sequence of a contig assembled from the
 CC entire cDNA insert in clone ssi.pk0022.al and a portion of the
 CC cDNA insert in clone sdp3c.pk004.n3 encoding the entire soybean
 CC acid triacylglycerol lipase (TAGL) (see AAY32309). The clones were
 CC isolated from soybean developing pod and seedling (5-10 day post
 CC germination) cDNA libraries. Novel acid and neutral TAGL polypeptides
 CC (see AAY32301-17) and polynucleotides (see AAY34950-66) from corn,
 CC Catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
 CC may be prepared recombinantly and used to raise antibodies, which
 CC are used for detecting the enzymes in situ in cells or in vitro in
 CC cell extracts. The polynucleotides may be used to create transgenic
 CC plants in which the TAGL levels are present at higher or lower levels
 CC than normal, or in cell types or developmental processes where they are
 CC not normally found. This would alter the level of triacylglycerol and
 CC cholesterol esters found in those cells. Accumulation of fatty acids
 CC with unusual structures may be a positive phenotype in plants used for
 CC foods. In addition, it may be desirable to eliminate expression of TAGL
 CC genes for certain applications. TAGL enzymes may also be useful for the
 CC processing of plant seed oils and for the development of novel seed
 CC oils. The TAGL enzymes can also be used as targets to facilitate the
 CC design and/or identification of inhibitors of those enzymes that may be
 CC useful as herbicides. This is desirable because inhibition of the
 CC activity of either of the enzymes could lead to an inhibition of plant
 CC growth. The polynucleotides also serve as a source of probes and
 CC primers, which are useful for genetic mapping, as markers for traits
 CC linked to those genes, and to isolate homologous sequences from other
 CC species.
 XX
 SQ Sequence 1718 BP; 492 A; 353 C; 344 G; 529 T; 0 other;
 Query Match 12.4%; Score 184.6; DB 21; Length 1718;
 Best Local Similarity 51.1%; Pred. No. 8.2e-40;
 Matches 489; Conservative 0; Mismatches 459; Indels 9; Gaps 2;
 QY 164 GGCCTCTGCGACGAGTCTCTCTGCGACATCGCGTACCCCTGCGACGACACAGCTTGA 223
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 310 GGCATATGTGCGCTTCTCTCTCATTTGTCATGATGATGATGATGATGATGATGATG 369
 QY 224 ACAAAAGATGGATTCCTTTTATCTCTTCAGCATATCCACATGGCAAAATAAAGCAGCA 283
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 370 ACTGATGATGGTTACATCTTGAGCCCTGCAAGGATCCCAAGAGGTGCGGTAAACAGT 429
 QY 284 GATAGT-----ACTGGCCCTCCAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGAC 337
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 430 GGGAGTGGGCAAGGAAGCAACAGTGGTTATACACATGGAGTTCTTCTAGATGGTATG 489
 QY 338 ACATGGTTTCATAACTCTGCTGAGCAATCATTGGGTATATCTTCTGATTAACGGTTTT 397
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 490 ACATGGCTTCTAAACCCAGCAGCAAGATCTCGCGTTGATTTTGTAGTATGATGATTT 549
 QY 398 GATGTTTGGATTGGGAATCTCGTGGCAACCGTTGGAGTAAAGGTCATTCAACCTTTTCT 457
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 550 GACGTGTGATTGCAAAACAAGAGGACCAAGATATAGTTCGCCGACACATCTCATTTGGAC 609

458 GTTCATGATAAGCTTTTCTGGGATGGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTA 517
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 610 CCTCTAGCCAGGCTTATTGGAATTGGTCTTGGGATGAACCTTGTCTCTCTATGATTTCCCT 669
 QY 518 GCANATGCTAGGCTATGCTATACAGTCACACAGTCCAAAATTCATATATGTGGGGCATTC 577
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 670 GCGGTGTTTAAATATATGTTTCAGCCAAACGGGCGAGAAGATCAATTTACGTTGGCCATTCA 729
 QY 578 CAGGAACTATATGTTTGGCGCTTTGACGATCCCCGAAATAGTAAAAATGATTAGC 637
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 730 TTGGGAACCTTTGGTAGCTTTGGCATCTTCTCGGAAGGAAATTTGGTTACCCAGCTGAA 789
 QY 638 TCTGCAGCACTTTCTTTGTCTTATTTCTTATCTTGATCATGTTAGTCTAGTTTGTCTC 697
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 790 TCAGCAGCTTGTGTAGCCCTATAGCCTTATTAAGCCACATGAATACAGCACTTGTGTT 849
 QY 698 AGAGCAGTCGCCATGCTTGTATCAGATGCTTGTACTATGGGAATTCACAGCTGAAC 757
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 850 GTTGCAACCAAGTCCCTTTGTTGGTAGATCACTACCCCTCTCGGCTAGCAGAAATTTAAT 909
 QY 758 TTCCGCTAGCGACATGGGCTTCAAAATAGTAGATTTCTTTGTGGATGGTGAACAGCTGGAT 817
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 910 CCAAAAGGTTAGCTTGTGATGCTTTCTCAAGTCTCTCTGTGCTCACCTCGGATAGAC 969
 QY 818 TGCACAATTTGCTATCTGCTGATTTACAGGGGAAAACCTGTTGCTTCAATACATCAAGGATT 877
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 970 TGTATGACTTGTGTGACTGCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029
 QY 878 GATTATTTTGGAGTATGAACCTCATCCATCATCGACAAAATAATCTGCACATCTTTT 937
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 1030 GATCTATTCTTGATGAATGAGCTCAGTCAACATCAACAAGAACATGGTGCACCTTGGCT 1089
 QY 938 CAGATGATCAGAAAGGCACCTTTGCGAAAGTATGACTATGGGTTATTGG--GAAACCTA 994
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 1090 CAGACTGTTAGACTTGGGCGTTGACAAAATTCATTTGTGAGACCCAGCATATAACATT 1149
 QY 995 AGCGCTACGGTCAITTTGCTCTCCCGCATTTGACCTAAGCAGCATACCAAGATCACTG 1054
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 1150 ATGCACTATGGAGAAATATTCTCTCCAATCTATAACCTTTTCCAACATCCCCCAGCATCTC 1209
 QY 1055 CCCATATGATGGGATATGAGGTCTTGTATGATGCTGGCTGATGTAACCGATGTTGAG 1111
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 1210 CCTCTCTTCACTAGCTATGGTGGAGAGATGCACCTTTTTCAGATGTCGCTGATGTTGAG 1266

RESULT 5
 AAI97406
 ID AAI97406 standard; cDNA; 802 BP.
 XX
 AC AAI97406;
 XX
 DT 13-NOV-2001 (first entry)
 XX
 DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3481.
 KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
 OS Homo sapiens.
 XX
 PN WO200166719-A1.
 XX
 PD 13-SEP-2001.
 XX
 PF 02-MAR-2001; 2001WO-JP01629.
 XX
 PR 07-MAR-2000; 2000JP-0159195.
 XX
 PA (CHIB-) CHIBA PREFECTURE.
 XX (HISM) HISAMITSU PHARM CO LTD.
 XX Nakagawara A;
 XX

QY	548	CAGTCCAAATTC	TATATGTGGGGCATT	CACAGGGAAC	TATATAGTGGTTTGGCGCTTTG	607
QY <td>548</td> <td>CAGTCCAAATTC</td> <td>TATATGTGGGGCATT</td> <td>CACAGGGAAC</td> <td>TATATAGTGGTTTGGCGCTTTG</td> <td>607</td>	548	CAGTCCAAATTC	TATATGTGGGGCATT	CACAGGGAAC	TATATAGTGGTTTGGCGCTTTG	607
DB	533	CAAGAACAAGT	GTATTATGTGGGTCT	CTCAGGCACC	ACTATAGTTTATAGCATTT	592
QY	608	AC---	GATGCCCGAATAGT	AAAAATGATTA	635	
DB	593	TCACAGATCCCT	GCAGCTGGCT	TAAGGATTA	623	
RESULT 8						
ABL57144						
ID	ABL57144	standard:	cDNA; 1411 BP.			
XX	AC	ABL57144;				
XX	XX	05-AUG-2002	(first entry)			
XX	XX	Kid goat pregastric	esterase gene.			
XX	XX	Esterase:	lipase; enzyme; goat; rennet; cheese; flavour; gene; ss.			
XX	XX	Capra hircus.				
XX	XX	Key	Location/Qualifiers			
XX	XX	CDS	68..1261			
XX	XX	FT	/*tag= a			
XX	XX	FT	/product= "Esterase"			
XX	XX	FT	68..124			
XX	XX	FT	/*tag= b			
XX	XX	FT	mat_peptide			
XX	XX	FT	125..1258			
XX	XX	FT	/*tag= c			
XX	XX	US6375947-B1.				
XX	XX	23-APR-2002.				
XX	XX	05-NOV-1998;	98US-0186489.			
XX	XX	05-NOV-1998;	98US-0186489.			
XX	XX	(INFL) INT FLAVORS &	FRAGRANCES INC.			
XX	XX	Bolen PL, Cihak PL,	Scharpf LG;			
XX	XX	WPI; 2002-424751/45.				
XX	XX	P-PSDB; ABB76189.				
XX	XX	An isolated kid goat	pregastric esterase free from other goat proteins			
XX	XX	is useful to prepare	enzyme modified cheese and to treat animals with			
XX	XX	an esterase or lipase	deficiency -			
XX	XX	Disclosure; Column	27-28; 35pp; English.			
XX	XX	The present sequence	is a nucleic acid encoding kid goat pregastric			
XX	XX	esterase (KPE; see	ABB76189). KPE free of other goat proteins can			
XX	XX	be produced by	purification from the kid gutlet or by recombinant			
XX	XX	expression in a non-	kid cell, especially an Escherichia coli,			
XX	XX	Saccharomyces	cerevisiae (NRRL Y-30030) or CHO cell (all claimed).			
XX	XX	KPE may be	produced with an N-terminal polyHis-esterokinase			
XX	XX	sequence, which	increases KPE expression. The KPE can be used			
XX	XX	to treat animals	with an esterase or lipase deficiency. The			
XX	XX	enzyme produces	a mixture of fatty acids when reacted with a dairy			
XX	XX	product, such as	lipolysed butter oil, milk, cheese or whey. It			
XX	XX	is therefore	useful in the production of enzyme modified cheeses			
XX	XX	as a substitute	for rennet, and in the creation of flavour in			
XX	XX	cheese foods or	any other food containing triglycerides. Microbial			
XX	XX	production will	allow the development of new Kosher and vegetarian			
XX	XX	food products.				
XX	XX	Sequence 1411 BP;	400 A; 315 C; 270 G; 426 T; 0 other;			

Query Match	6.1%	Score 91.2	DB 24	Length 1411
Best Local Similarity	54.2%	Pred. No. 2.3e-14		
Matches 230	Conservative 0	Mismatches 188	Indels 6	Gaps
QY 194	GGCTACCCCTGCACCGACACACGTTTAAACAAAGATGGATTCTCTTTATCTCTTCAG 253			
Db 191	GGTACCCCAAGTCAGATGCATAAAGTTATACTGCAGATGGCTATATCCTTCAGGCTAT 250			
QY 254	CATATCCCAACATGGCAAAAATAAAGCAGCAGAT---AGTACTGGCCCTCCAGTTTTCTTT 310			
Db 251	CGGATTCTTCATGGAAAGAATGATGCTAATCATTTAGGTCACAGACCTGTTGTTTCTG 310			
QY 311	CACATATGGCTTTTTCAGGAGGAGACACATGGTTCAATAACTCTGCTGAGCAATCACTT 370			
Db 311	CAGCATGGTCTTCTTCGCCCTCAGCTACAAACTGGATTTCACAGCTTCCCACACACAGCCTG 370			
QY 371	GGGTATATCTCTCTGATACAGGTTTTTGATGTTTGGATTGGGAATGTCCTGCGAAGCGT 430			
Db 371	GGCTTCTCTGCGAGATGCTGGTTATGACGTGTGCTGGGGAACAGCAGAGGAACACT 430			
QY 431	TGGAGTAAAGTCAATCAACCTTTTCTGTTCATGATAAAGCTTTCTTGGGATTTGGAGCTGG 490			
Db 431	TGGGCCCGAGCAATTTATATCATTCACCAGACTCCCTCGAATCTGGGCTTTCAGCTTT 490			
QY 491	CAGAGTTAGCTGAATATGACCTTTTAGCAATGCTAGGCTATGTGT---ATACAGTCACA 547			
Db 491	GATGAAATGGCTGAATATGACCTTCCATCTACAATTGATTTTCATCTTAAAGAGAACAGGA 550			
QY 548	CAGTCCAAAATTCATATGTTGGGGCATTCACAGGGAACCTATATATGGGTTTGGCGGCTTTG 607			
Db 551	CAGAAGAAGCTACACTATGTGGCCATTTCCCAAGGCACCACCATTTGGTTTTTGTGCGCTTT 610			
QY 608	ACGA 611			
Db 611	TCTA 614			
RESULT 9				
ABL57145				
ID	ABL57145 standard; DNA; 8324 BP.			
XX	AC ABL57145;			
XX	05-AUG-2002 (first entry)			
XX	Yeast YE-1 expression vector.			
DE	Vector; YE-1; esterase; lipase; goat; cheese; flavour; ds.			
XX	Saccharomyces cerevisiae.			
XX	US6375947-B1.			
PN	23-APR-2002.			
PD	05-NOV-1998; 98US-0186489.			
XX	05-NOV-1998; 98US-0186489.			
XX	{INFL } INT FLAVORS & FRAGRANCES INC.			
PR	Bolen PL, Cihak PL, Scharpf LG;			
XX	WPI; 2002-424751/45.			
XX	An isolated kid goat pregastric esterase free from other goat proteins			
PT	is useful to prepare enzyme modified cheese and to treat animals with			
PT	an esterase or lipase deficiency			
XX	Disclosure; Column 29-38; 35pp; English.			
PS	The present sequence is that of yeast YE-1 expression vector.			
XX	The vector comprising a nucleic acid (see ABL57144) encoding kid			

CC	goat pregastric esterase (kPE, see AB76189) can be used to
CC	transform a host cell, such as Saccharomyces cerevisiae, for
CC	recombinant production of kPE. The recombinant kPE can be used
CC	to treat animals with an esterase or lipase deficiency. The enzyme
CC	produces a mixture of fatty acids when reacted with a dairy product
CC	such as lipolyzed butter oil, milk, cheese or whey. It is
CC	therefore useful in the production of enzyme modified cheeses as a
CC	substitute for rennet, and in the creation of flavour in cheese
CC	foods or any other food containing triglycerides. Microbial
CC	production will allow the development of new kosher and vegetarian
CC	food products.
XX	
SQ	Sequence 8324 BP; 2394 A; 1884 C; 1703 G; 2343 T; 0 other;
Query Match	6.1%; Score 91.2; DB 24; Length 8324;
Best Local Similarity	54.2%; Pred. No. 5,7e-14;
Matches 230; Conservative	0; Mismatches 188; Indels 6; Gaps
QY	194 GGCTACCCCTGCACGAGCACAACGTGTAACAAAGAATGGATTCTTTTATCTTCACG 253
Dd	
Dd	1598 GGCTAACCAAGTGAGATGCATAAAGTTATAACTGCAGATGGCTATATCCTTCAGGTCTAT 1657
QY	254 CATATCCCACATGCCAAAATAAAGCAGCAGAT---AGTACTGGCCCCCTCCAGTTTTCTT 310
Dd	
Dd	1658 CGGATTCTCATGGAAGAATGATGCATAATTTAGCTCAGAGACCTGTTGCTGTTCTG 1717
QY	311 CAACATGGTCTTTTCCAGGAGGAGACACATGGTTCATAAACTCTGCTGAGCAATCACTT 370
Dd	
Dd	1718 CAGCATGGTCTCTCGCTCAGCTCAAACACTGGATTTCCAACCTTCCCAACACAGCGCTG 1777
QY	371 GGGTATATCTTCTCTGATACAGGTTTTGATGTTTGGATGGGAATGTCCTGGGAACCGT 430
Dd	
Dd	1778 GGCTTCTCTCGCACATGCTGTTATGACGTGTGCTGGGGAACAGCAGAGGAACACT 1837
QY	431 TGAGATAAGGTCAATCAACCTTTTCTGTTATGATATAGCTTTTCTGGGATGGAGCTGG 490
Dd	
Dd	1838 TGGGCCCCAGGAACATTTATACTATTCCACAGACTCCCTGTAATTTGGGCTTTCAGCTTT 1897
QY	491 CAGAGGTACGCTCAANTATGACCTTTTAGCAATGCTAGGCTATGTGT---ATACAGTCACA 547
Dd	
Dd	1898 GATGAATGGCTGAATATGACCTTCCATCTACANTGATTTTCATCTTAAAGAGAACAGGA 1957
QY	548 CAGTCCAAAAATCTATATGTGGGCATTCACAGGAACTATATGGGTTTGGCGGCTTTG 607
Dd	
Dd	1958 CAGAAGAAGCTACACTATGTGGCCATTCCCAAGSCACCACCATTTGGTTTGTGCGCTTT 2017
QY	508 ACGA 611
Dd	
Dd	2018 TCTA 2021
RESULT 10	
ABK86570	
ID	ABK86570 standard; cdNA; 1116 BP.
XX	
AC	ABK86570;
XX	
DT	24-SEP-2002 (first entry)
XX	
DE	Human cDNA encoding lysosomal acid lipase #2.
XX	
KW	Human; ss: gene; lysosomal acid lipase; lipid malabsorption illness;
KW	cystic fibrosis; alcoholism; heart disease; heart attack; injury;
KW	Wolman disease; cholesterol ester storage disease; brain injury;
KW	mood disorder; anxiety disorder; thought disorder; volition disorder;
KW	sleep disorder; neurogenic disorder; myopathic disorder; COPD;
KW	obesity; cancer; chronic obstructive pulmonary disease; diabetes;
KW	cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
KW	anorexia; osteoarthritis; central nervous system disorder;
KW	peripheral nervous system disorder.
XX	
OS	Homo sapiens.
XX	


```
CC treatment of a disease such as lipid malabsorption illness,
CC cystic fibrosis, alcoholism, heart disease, heart attack,
CC Wolman disease, cholesterol ester storage disease, brain injury,
CC mood disorder, anxiety disorder, thought disorder, volition disorder,
CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
CC chronic obstructive pulmonary disease (COPD), diabetes,
CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
CC anorexia, osteoarthritis, a central nervous system disorder and
CC a peripheral nervous system disorder. The present sequence is the
CC cDNA encoding human lysosomal lipase #1.
XX
SQ Sequence 1210 BP; 338 A; 264 C; 296 G; 312 T; 0 other;

Query Match          6.0%; Score 89.4; DB 24; Length 1210;
Best Local Similarity 53.6%; Pred. No. 6.5e-14;
Matches 233; Conservative 0; Mismatches 196; Indels 6; Gaps 2;

QY 183 TCCTGCCACTGGCTACCCCTGCACCGAGCACACAGTGTGAACAAAGATGGATTCCTTT 242
DB 83 TCCAAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTCAAGATGGTATATCC 142
QY 243 TATCTCTTCAGCATATCCACATGCGCAAAATAAGCAGCAGATA--GTACTGGCCCTC 299
DB 143 TTTCTGTTAAACAGGATTCCTCGAGCCCTAGTGCACCTTAAGAAGACAGGTTCCAGGCCGTG 202
QY 300 CAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGACACATGTTTCATAAATCTGCTG 359
DB 203 TGGTGTCTAGCAGCATGGCTAGTTGGAGGTGCTAGCACTGGATTTCCAACTGCCCA 262
QY 360 AGCAATCACTTGGGTATATCTCTGCTGATACGGTTTGTGATGTTGGATGGGAATGCC 419
DB 263 ACAATAGGCTGGGCTTCATCTGGCAGATGCTGTTTTCAGCTGTGGATGGGAACACGA 322
QY 420 GTGGAAACGCTTGGAGTAAAGTCTATCAACCTTTCTGTTTCATGATAAGCTTTTCTGGG 479
DB 323 GGGGAACGCTGCTGCGAAACACAGACACTCTCCATAGACCACAGATGAGTCTGGG 382
QY 480 ATTGAGCTGGCAAGAGTTAGTGAATATGACCTTTTACCAATGCTAGGCT---ATGTGT 536
DB 383 CTTTCAGTTATGATGAGATGGCTAGGTTTTCAGCTTCTCTGCAAGTATAAATTTATTTCG 442
QY 537 ATACAGTCACACAGTCCAAATCTATATCTGCGGCATTCACAGGAACTATATGGGTT 596
DB 443 AGAAACGGCGCCAGAAAGATCTATTATGTCGGCTATTTCACAGGGCCACCACATGGGCT 502
QY 597 TGGCGGCTTTGACGA 611
DB 503 TTATTGCATTTTCCA 517

RESULT 12
AAF45132
ID AAF45132 standard; cDNA; 1269 BP.
XX
AC AAF45132;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 ORF.
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS Homo sapiens.
XX
PN WO200077239-A2.
XX
PD 21-DEC-2000.
XX
```

```
PF 24-MAY-2000; 2000WO-US14858.
XX
PR 14-JUN-1999; 99US-0333159.
XX (MILL-) MILLENNIUM PHARM INC.
XX
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX WPI; 2001-032313/04.
DR P-PSDB; AAB66065.
XX
PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease.
XX
PS Claim 1; Fig 6; 359pp; English.
XX
CC The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.
XX
SQ Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

Query Match          6.0%; Score 89.4; DB 22; Length 1269;
Best Local Similarity 53.6%; Pred. No. 6.6e-14;
Matches 233; Conservative 0; Mismatches 196; Indels 6; Gaps 2;

QY 183 TCCTGCCACTGGCTACCCCTGCACCGAGCACACAGTGTGAACAAAGATGGATTCCTTT 242
DB 158 TCCAAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTCAAGATGGTATATCC 217
QY 243 TATCTCTTCACATATCCACATGCGCAAAATAAGCAGCAGATA--GTACTGGCCCTC 299
DB 218 TTTCTGTTAAACAGGATTCCTCGAGGCCCTAGTGCACCTTAAGAAGACACAGGTTCCAGGCCGTG 277
QY 300 CAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGACACATGTTTCATAAATCTGCTG 359
DB 278 TGGTGTACTGCAGCAUGGCCCTAGTTGGAGGTGCTAGCACTGGATTTCCAACTGCCCA 337
QY 360 AGCAATCACTTGGGTATATCCTTCTGATAACGGTTTTCATGTTTGGATTGGGAATGTCC 419
DB 338 ACAATAGCTGGGCTTCATCTGCGCAGATGCTGGTTTTCAGCTGTGGATGGGAACAGCA 397
QY 420 GTGGAACGCTTGGAGTAAAGTTCATCAACCTTTTCTGTTTCATCATGATAAGCTTTTCTGGG 479
DB 398 GGGGAACGCTGGTCTCGAAACACAAAGACTCTCCATAGACCAAGATGAGTTCTCGG 457
QY 480 ATTGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTTCAGCAATGCTAGGCT---ATGTGT 536
DB 458 CTTTCAGTTATGATGAGATGGCTAGGTTTTCAGCTTCTGTCAGTATAAATTTATTTCG 517
QY 537 ATACAGTCACACAGTCCAAATCTTATATGTTGGGCAATTCACAGGAACTATAATGGGTT 596
DB 518 AGAAACGGCGCCAGAAAGATCTATTATGTCGGCTATTTCACAGGGCCACCACATGGGCT 577
QY 597 TGGCGGCTTTGACGA 611
DB 578 TTATTGCATTTTCCA 592

RESULT 13
AAD19220
ID AAD19220 standard; DNA; 1384 BP.
```



```
Db 403 TGGTGTACTGCAGCATGGCCTAGTGGAGGTGCTAGCAACTGGATTCCAACTGCGCA 462
QY 360 AGCAATCACTTGGGTATATCCTTGCTGATAAGGGTTTGGATCTTTGGATTTGGGAATGTCC 419
Db 463 ACAATAGCCTGGGCTTCATCTGGCAGATGCTGGTGTGGATGGGAAACAGCA 522
QY 420 GTGGAACGGTTGGAGTAAAGTCAATCAACCTTTCTGTTTCATGATAAGCTTTTCTGGG 479
Db 523 GGGGAAACGGCTGGTCTCGAAACACAAGACACTCTCCATAGACCAAGATGAGTTCTGGG 582
QY 480 ATTGAGCTGGCAAGCTTAGCTGAATATGACCTTTTAGCAATGCTAGGCT---ATGTGT 536
Db 583 CTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCTCCTGCAGTGATAAACTTTATTTTGC 642
QY 537 ATACAGTCACACAGTCCTCAAAATTCATATATGTGGGGCATTCACAGGGAACATATAATGGGT 596
Db 643 AGAAACGGCCAGGAAAGATCTATTATGTGGCTATTTCACAGGGCACACCATTGGGCT 702
QY 597 TGGCGGCTTTGACGA 611
Db 703 TTATTGCATTTTCCA 717
```

Search completed: December 30, 2002, 17:23:18
Job time : 291 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2002, 17:13:26 ; Search time 2226 Seconds
(without alignments)
10789.716 Million cell updates/sec

Title: US-09-699-652A-13

Perfect score: 1483

Sequence: 1 gcaagagatcacagcgggc.....ctctcaaaaaaaaaaaaaa 1483

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	549.6	37.1	717	10 BE443524	BE443524 WHE1115_C
2	464.6	31.3	623	13 BM324154	BM324154 PIC1_24_G
3	403	27.2	668	13 BI075779	BI075779 IPL_21_B0
4	360	24.3	491	10 BE367129	BE367129 PIC1_43_D1
5	357.4	24.1	478	9 AU088737	AU088737 AU088737
6	341.6	23.0	631	13 BM037885	BM037885 S113f11 S

7	327.6	22.1	828	14	BQ122157	BQ122157 EST607733
8	327.4	22.1	441	13	BM501471	BM501471 PAC000000
9	324.2	21.9	365	9	AU161807	AU161807 AU161807
10	324.2	21.9	572	13	BM326139	BM326139 PIC1_66_B
11	316.4	21.3	476	10	BE367231	BE367231 PIC1_43_D1
12	294.4	19.9	647	13	BM328649	BM328649 PIC1_24_G
13	285.8	19.3	323	9	AU161802	AU161802 AU161802
14	285.2	19.2	646	13	BI075485	BI075485 IPL_21_B0
15	281	18.9	521	9	AU088738	AU088738 AU088738
16	280.4	18.9	420	9	AJ474806	AJ474806 AJ474806
17	277.6	18.7	449	9	AU174368	AU174368 AU174368
18	263.8	17.8	635	14	BQ122609	BQ122609 EST608185
19	263.2	17.7	600	10	BE191250	BE191250 sn94a09.y
20	253	17.1	659	13	BI405491	BI405491 13BE11 Ma
21	245.6	16.6	588	13	BM331443	BM331443 PIC1_66_B
22	240.4	16.2	549	13	BM527928	BM527928 sal52810
23	231.2	15.6	394	10	BE356956	BE356956 DGI_145_F
24	227.2	15.3	501	13	BI784642	BI784642 sal57604
25	225.2	15.2	492	9	AL380410	AL380410 MCB52C08
26	215.2	14.5	422	14	BQ756230	BQ756230 EB6m06_SQ
27	213.6	14.4	536	13	BI321617	BI321617 sal15601
28	212	14.3	684	14	BQ407284	BQ407284 GA_E010
29	210.6	14.2	630	14	BQ415924	BQ415924 GA_E4010
30	209.6	14.1	600	10	AV559339	AV559339 AV559339
31	204.6	13.8	563	10	AW277778	AW277778 sf86e03.y
32	204.6	13.8	625	12	BE822354	BE822354 GM700017A
33	203	13.7	702	12	BF479168	BF479168 L48-282T
34	202.2	13.6	469	12	BF325456	BF325456 su33a06.y
35	201.4	13.6	650	14	BQ124957	BQ124957 EST610533
36	199.4	13.4	591	14	BQ743009	BQ743009 sqs58909
37	197.2	13.3	624	12	BE822717	BE822717 GM700018A
38	195.6	13.2	329	12	BG313685	BG313685 WHE2057_C
39	195.2	13.2	461	12	BE800757	BE800757 sq97h02.y
40	193.6	13.1	558	13	BI427293	BI427293 sab78e08
41	190.4	12.8	416	10	AW830986	AW830986 sm31d03.y
42	189.6	12.8	573	10	BE318901	BE318901 NF004005L
43	188.8	12.7	514	14	BQ122585	BQ122585 EST608161
44	185.6	12.5	446	12	BF716146	BF716146 saa15606
45	185.4	12.5	340	9	AU223277	AU223277 AU223277

ALIGNMENTS

RESULT 1
BE443524

LOCUS

DEFINITION

BE443524

WHE1115_C05_F092S Wheat etiolated seedling root normalized cDNA

Library Triticum aestivum cDNA clone WHE1115_C05_F09, mRNA

sequence..

ACCESSION

BE443524

VERSION

BE443524.1

KEYWORDS

EST..

SOURCE

bread wheat.

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 717)

AUTHORS

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han

, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Nguyen, H.T.,

Rausch, C.J., Seaton, C.L., Tong, J.C. and Zhang, D.

The structure and function of the expressed portion of the wheat

genomes - Normalized root cDNA library

Unpublished (2000)

JOURNAL

Contact: Olin Anderson

COMMENT

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20
Seq primer: Stragatene SK primer.

FEATURES

source
1..717
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE115_C05_F09"
/clone_lib="Wheat etiolated seedling root normalized cDNA library"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli DH10B"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the T3 Close lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give pBluescript phagemids before normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
182 a 137 c 163 g 235 t
BASE COUNT
ORIGIN
Query Match 37.1%; Score 549.6; DB 10; Length 717;
Best Local Similarity 86.7%; Pred. No. 9.3e-118;
Matches 617; Conservative 0; Mismatches 94; Indels 1; Gaps 1;
QY 223 AACAAAGATGGATTCCTTTATCTCTCCAGCATATCCACATGGCAAAATAAAGCAGC 282
DB 7 AAACAGAGATGGCTTTCTTTCTCTCCAGCATATTCACATGGCAAAATGGAGTTCC 66
QY 283 AGATAGTACTGCCCTCCAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGACACATG 342
DB 67 AGATAATACTGGACCTCCAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGATACATG 126
QY 343 GTTCATAACTCTGCTAGCAATCACTTGGGTATATCTTGTGTGATAACGGTTTGTATCT 402
DB 127 GTTCATAAATCTGCTAGCAGCTGGTGGATATATCTTGTGTATATGGGTTTCGATCT 186
QY 403 TTGGATTGGGAATGTCCGGTGAACCGGTTGGAGTAAAGTCAATTCACACCTTTTCTGTCCA 462
DB 187 TTGGATTGGAATGTTCGTGGAACACGTTGGAGTAAAGGCCATTCACATTCACATGTGCA 246
QY 463 TGATAGCTTTTCTGGGATGGAGCTGGCAAGAGTTAGCTGATATGACCTTTTAGCAAT 522
DB 247 TGATAAGCTTTTCTGGGATGGAGCTGGCAAGAGCTTGTGTAATATGATCTTCTGGCAAT 306
QY 523 GCTAGGCTATGTGTATACAGTACACAGTCCAAATTCATATGTGGGCAATTCACAGGG 582
DB 307 GTTAAGCTAGGTATACAGTATAGGAGTCCAAATTTTGTATGTGGGCAATTCACAGGG 366
QY 583 AACTATAATGGGTTGGCGGCTTTGACGATGCCGGAATAGTAAATGATTAGTCTCTGC 642
DB 367 AACTATTATGGGTTTGGCTGCTTTTACATGCTGGAATACAAAGATGATTAGCGCTCC 426
QY 643 AGCACTTCTTTGTCCTATTTCTTAATCTGATCATGTAGTGTAGTTTGTCTCAGACG 702
DB 427 TGCATCTCTTTGTCCAAATTTCTTACCTTCATCATGTAGTGTAGTTTGTCTCAGACG 486
QY 703 AGTCGCATCATCTGTATGATGCTGTTACTATGGGAATTCACCACTGAACCTCCG 762
DB 487 AGTGGCATGCATCTTGACACAGATGCTTCTTACCATGGGCTCCATCAGCTGAACCTCCG 546
QY 763 TAGCGACATGGGGTTTCAAATAGTAGATTCTTTTGTGCGGATGTGAACACGTGGATTGCAA 922

Db 547 GAGCGATATGGGAGTTCAAATAGTAGATTCTATATGCGATGATGACATGTGGACTGCAA 606
QY 823 CAATTTGCTATCTGCGATTACAGGGAAACTGTTGCTTCATACATCAAGGATTGATTA 882
Db 607 CGATTTGCTGTCTTC-ATAACAGGGGAAATGTTGCTTCATGATCAGGATTGACCA 665
QY 883 TTATTTGGAGTATGAACCTCATCCATCGACAAAAAATCTGCACCATCTT 934
Db 666 TTACTTGGAGTAGAACCTCATCCATCTCACTAAAAAATCTGCATCTT 717
RESULT 2
BM324154
LOCUS
DEFINITION
BM324154 623 bp mRNA linear EST 04-JAN-2002
bicolor cDNA, mRNA sequence.
ACCESSION
BM324154
VERSION
BM324154.1 GI:18062592
KEYWORDS
EST.
SOURCE
sorghum.
ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 623)
AUTHORS
Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
Sudman,M. and Pratt,L.H.
TITLE
An EST database from Sorghum: plants infected with a compatible
pathogen
JOURNAL
Unpublished (2002)
COMMENT
Contact: Cordonnier-pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector, and regions
below Phred quality 16. The threshold for highest quality sequence
is 20. Three-prime sequences, which are obtained with PolyMix or
T7 sequencing primer, are presented as the reverse complement.
Seq primer: JEN REV
High quality sequence stop: 573
POLYA=No.
FEATURES
Location/Qualifiers
1..623
/organism="Sorghum bicolor"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone_lib="Pathogen-infected compatible 1 (PIC1)"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with
Colletotrichum graminicola"
/note="Vector: pBluescript II SK(-) from Lambda Zap II;
Site_1: XhoI; Site_2: EcoRI; Four-week-old sorghum
seedlings were sprayed with spore suspension prepared from
3-week-old FRM421, a sorghum isolate of the anthracnose
pathogen Colletotrichum graminicola. Inoculated plants
were kept in a 25 C dark growth chamber with 100% relative
humidity for 24 hr, followed by 12/12 hr of light/dark
cycle at 25 C with 90% relative humidity for another 24
hr. All leaves were harvested and quick frozen with liquid
nitrogen and stored in a -80 C freezer. The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision.
WARNING: While most or all ESTs are expected to derive
from the host plant, no effort was made to eliminate ESTs
deriving from the pathogen."
158 a 116 c 139 g 210 t
BASE COUNT
ORIGIN
Query Match 31.3%; Score 464.6; DB 13; Length 623;

Best Local Similarity 84.1%; Pred. No. 6.4e-98;
Matches 524; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```
QY 267 GCAAAATAAAGCAGACAGATAGTGGCCCTCCAGTTTCTTCAACATGCTCTTTTC 326
Db 1 GCACAGGGGAATACAGATAATGCTGGACCTCCAGTTTTCTTCAGCATGCTCTTTTC 60
QY 327 AGGAGGAGACACATGGTTTCATAACTCTGCTGAGCAATCACTTGGGTATATCTTTGCTG 386
Db 61 AGGGTGGAGATACATGGTTTCATAACTCAATGAACAATCACTTGGATATATCTTTGCTG 120
QY 387 ATAACGGTTTTGATGTTGGATGGAATGCTCGGGAACGGTTGGAGTAAGGTCAAT 446
Db 121 ATAATGGTTTTGATGTTGGATGGAATGCTCGGGAACGGTTGGAGTAAGGCCACT 180
QY 447 CAACCTTTCTGTCATGATAAGCTTTCTGGGATTGGAGCTGGCAAGAGTTAGTGAAT 506
Db 181 CTACTCTCTGTCATGATAAGCTTTCTGGGAATGGAGTTGGCAAGACCTTGGCTGAAT 240
QY 507 ATGACCTTTTAGCAATGCTAGGCTATGTATACAGTACACAGTCCAAATTCATATG 566
Db 241 ACGATGTTTTGGCAATGTTAAGCTATGTATATACAATTAACAGTCCAAATTTTCATATG 300
QY 567 TGGGGCATTCACAGGAACATAAATGGGTTTGGCGCTTTGACGATGCCGAAATAGTAA 626
Db 301 TGGGACATTCACAGGGAACATATCATGGGTCTGGCTGCGTTTACAATGCCCTGAATAGTAA 360
QY 627 AAATGATTAGCTCTCAGACACTCTTTGTCCTATTTCTTATCTTGATCATGTAGTGCTA 586
Db 361 AAATGATAAGCTCTGCTGTCCTTTGTCCTATTTCTTACCTTGATCATCACTGCTGCTA 420
QY 687 GTTTTGTCTCAGAGCAGTCGCATGCTTTGATCAGATGCTTTGATCTAGTGGGAATTC 746
Db 421 GTTTTGTCTTAGAGCAGTTGCCATGCTTTGACCATCTTGACCATGCTTTGCTATGGGCATCC 480
QY 747 ACCAGTGAACCTTCGTAGCAGACATGGGGTTTCAATAGTATGATCTTTTGTGCGATGGTG 806
Db 481 ATCAGTTGAACCTTCGGAGTGATATGGGTGTCAGATTTTAGATTGCTGCTGATGATG 540
QY 807 AACAGTGGATTGCAACAATTTGCTATGCGATTCAGCGGGAACACTCTTCTTCAATA 866
Db 541 ACATTTGACATGCAACGATCTGTTATCTTCAATACAGGTCAAAACTGTTGTTTCAAT 600
QY 867 CATCAAGGATTGATTATTTTGG 889
Db 601 CATCTCGGATTGACTATTATTG 623
```

RESULT 3
BI075779
LOCUS
DEFINITION
IP1_21_B03_b1_A002 Immature pannicle 1 (IP1) Sorghum bicolor cDNA,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BI075779
BI075779.1 GI:14514436
EST.
sorghum.
Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
AUTHORS
Klein,R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt
L.H.

TITLE
JOURNAL
COMMENT
An EST database from Sorghum: developing preanthesis pannicles
Unpublished (2001)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with PolyMix or T7
sequencing primer, are presented as the reverse complement.

seq primer: JEN REV
High quality sequence stop: 572
POLYA-No.

FEATURES
source
1. 668
/organism="Sorghum bicolor"
/cultivar="Brx623"
/db_xref="taxon:4558"
/clone_lib="Immature pannicle 1 (IP1)"
/note="Organ: Developing preanthesis pannicles; Vector:
pBluescript II SK(-) from Lambda Zap II; Site: 1; XhoI;
Site: 2; EcoRI; The library was made from poly-A RNA in the
cloning vector lambda Zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT 145 a 176 c 175 g 171 t 1 others
ORIGIN

Query Match 27.2%; Score 403; DB 13; Length 668;
Best Local Similarity 78.9%; Pred. No. 1.5e-83;
Matches 494; Conservative 0; Mismatches 126; Indels 6; Gaps 1;
QY 23 GGTGTTGGCGATGGCGGGCCAGCCCGGAGGAGCGCTCCCTCTGATCTCTCCTC 82
Db 42 GCGGTGGCGATGGTGGCCCGAGGAGCGCTCGCGGGCCCGCGCTCTCTGTGCTC 101
QY 83 GTGCTCTCTGCTGGCGTGCATCTCTCGGAGGCTCCCAAGCGCGCCCGCTCCGC 142
Db 102 GTGCTCTCTCTCTCTCGCGGTGAGCGCCGCGCATCCCGGCCACACAGCGCTCGC 161
QY 143 CGGCTCGGCTCC-----GGCTCGGCGGCTCTCGGACCAGCTGCTCTGCCACTCGGC 196
Db 162 CGGCTCTCCCGCTCGCGCGCGCGCGCTCTGCCAGCAGCTGCTCTGCCCGAGGC 221
QY 197 TACCCTCGCACGACACACCTTTGAAACAAAAGATGATCTTTATCTCTTCAGCAT 256
Db 222 TACCCTCGCACGACACACCTTTCAACGGATGATGCTCTCTTTGCTCTTCAGCAT 281
QY 257 ATCCACATGGCAAAATAAAGCAGCAGATAGTACTGGCCCTCCAGATTTTCTTCAACAT 316
Db 282 ATTCACATGGCAAAATGGAATACAGATAATGCTGGACCTCCAGATTTTCTTCAGCAT 341
QY 317 GGTCTTTTTCAGGAGGAGACACATGTTCTATAAACTCTGCTGACCAATCACTTGGGTAT 376
Db 342 GGTCTTTTTCAGGAGGAGATACATGGTTCTATAAACTCCAATGAACAATCACTTGGATAT 401
QY 377 ATCTTGTGATAACCGTTTGTGATTTGGATTGGGAATGTCCTGGAAACGGCTTGGAGT 436
Db 402 ATCTTGTGATAATGTTTGTGATTTGGATTGGAAATGTCCTGGACACAGTTGGAGT 461
QY 437 AAAGTCATTTCAACCTTTTCTGTCATGATAAGCTTTTCTGGGATTGGAGCTGGCAAGAG 496
Db 462 AAAGGCCACTCTACTCTCTCTCTTCATGATAAGCTTTTCTGGGAATGGAGTTGGCAAGAC 521
QY 497 TTAGCTGAATATGACCTTTTAGCAATGCTAGCTATGTATATACAGTACACAGTCCAAA 556
Db 522 CTGTGTAATAGCATGTTTGTGCAATGTTAAGCTATGTATATACAAATTACACAGTCCAGA 581
QY 557 ATTCTATATGTGGGCATTTCACAGGGAACATAAATGGTTTGGCGCTTTTGACGATGCC 616
Db 582 ATTCTATATGTGGGACATTCACAGGGAACATATCATGGGTCTGGCTGCGTTTACNATGCT 641
QY 617 GAAATAGTAAAAATGATTAGCTCTGC 642
Db 642 GAAATAGTAAAAATGATAGCTCTGC 667

RESULT 4
BE367129
LOCUS
BE367129 491 bp mRNA linear EST 20-JUL-2000

DEFINITION P11_43_D12.bl_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE367129
VERSION BE367129.1 GI:9308686
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 491)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
L.H.

TITLE An EST database from Sorghum: pathogen-induced plants

JOURNAL Unpublished (2000)

COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.

Seq primer: JEN REV

High quality sequence stop: 458

POLYA=No.

FEATURES Location/Qualifiers

source 1..491

/organism="Sorghum bicolor"

/db_xref="taxon:4558"

/clone_lib="Pathogen induced 1 (P11)"

/note="Organ: Anthracnose-infected leaves from

two-week-old sorghum plants 48 hr after inoculation;

Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;

Site_2: EcoRI; Two-week-old sorghum plants (BTX 623

cultivar) were infected with pathogen (isolate FRM421 of

Colletotrichum graminicola, which is a sorghum isolate).

RNA was prepared from infected leaves harvested from 45

seedlings 48 hours after inoculation. Note: young

seedlings (2 weeks old) exhibit juvenile resistant

reaction, which is an incompatible interaction. As they

grow older (4 weeks or older), plants resume susceptibility

to anthracnose disease. The library was made from poly-A

RNA in the cloning vector Lambda Zap II. Clones to be

sequenced were prepared by mass excision. WARNING: While

most or all ESTs are expected to derive from the host

plant, no effort was made to eliminate ESTs deriving from

the pathogen."

BASE COUNT 131 a 104 c 122 g 134 t

ORIGIN

Query Match 24.3%; Score 360; DB 10; Length 491;

Best Local Similarity 83.6%; Pred. No. 1.7e-73;

Matches 408; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 706 CGCCATGCATCTTGATCAGATCGTTGTACTATGGGATTCACGAGCTCCGCTAG 765

DB 4 CGAGTGCATCTTGACAGATCGTTGTGCTATGGGCATCCATCAGTTGAACCTCCGGAG 63

QY 766 CGACATGGGGTTCAANTAGTAGATCTTTGTGCGATGCTGAACACGTTGGATTGCAACAA 825

DB 64 TCATATGGGTGTCAGATTTTATAGTCGCTGTGATGATGACATTTGGAGTCAACGA 123

QY 826 TTTGCTATCTCGGATTACAGGGGAAACATGTTGCTTCAATACATCAAGATTTGATTATA 885

DB 124 TCTGTTATCTTCAATAACAGGTCAAACATGTTGTTTCAATTCATTCGAGTACTATTA 183

QY 886 TTTGGAGTATGAACCTCATCCATCATCCACAAAAATCTGCACATCTTTTCAGATGAT 945

DB 184 TTTGGAGTATGAACCCCATCCGCTCATCAACAAAAATTTGGGCGCATCTTTTCCAGATGAT 243

QY 946 CAGAAAGGCACCTTCCAAAGTATGACTATGGTTATTGGGAAACCTAAGCGCTACGG 1005
DB 244 CAGAAAGGCAGTTTCGCAAGTATGACTATGGTTGGGAAACATAAGGCACCTACGG 303
QY 1006 TCATTTGGTCTCTCCGATTTTGACCTTAAGCAGCATACCAGATCACTGCCCATATGAT 1065
DB 304 CCACGGCGCATCTCTCCCTCATTCGATTAAGCAGCATACCGGAATCACTGCCAATATGAT 363
QY 1066 GGGATATGGAGGTCTTGATGCATTGGCTGATGAACCGATGTTCCAGCGTACTACTACAGA 1125
DB 364 GGGTATGGAGGCTTGTATGCACCTGGCTGACGTAAACAGATGTTGAGCGCACCATCAAGA 423
QY 1126 GCTGGGATCTACACAGAACTTCTGTACATTTGTCACCTATGCGCATATGATTTTGTAT 1185
DB 424 GTTGAGATCCACGCGCAGAGCTGCTGTACATTTGTTGACTATGCGCACATTTGATTCATCAT 483
QY 1186 GAGCGTGA 1193
DB 484 GAGCGTGA 491

RESULT 5

AU088737

LOCUS

DEFINITION

AU088737 Rice shoot Oryza sativa

clone S5305, mRNA sequence.

ACCESSION

AU088737

VERSION

AU088737.1

KEYWORDS

EST.

SOURCE

Oryza sativa (japonica cultivar-group).

ORGANISM

Oryza sativa (japonica cultivar-group).

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

Sasaki,T. and Yamamoto,K.

TITLE

Rice cDNA from etiolated shoot (2000)

JOURNAL

Unpublished (2000)

COMMENT

Contact: Takuji Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@nri.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/

PROJECT = 'RGP'

FEATURES

Location/Qualifiers

source 1..478

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="S5305"

/clone_lib="Rice shoot"

/dev_stage="Etiolated shoot (8 days old)"

/note="Etiolated shoot (8 days old)"

BASE COUNT 98 a 148 c 120 g 110 t 2 others

ORIGIN

Query Match 24.1%; Score 357.4; DB 9; Length 478;

Best Local Similarity 94.6%; Pred. No. 6.8e-73;

Matches 434; Conservative 0; Mismatches 18; Indels 7; Gaps 6;

QY 10 CACAGCGCGCGGCGTTGGCGATGGCGATCGCGGCCACGCCCGGAGGAGCGCTCCC 69

DB 24 CAAGCGCGCGGCGGCTTGGCGATGGCGATCGCGGCCACG-CCCCGAGGAGCGCTCCC 82

QY 70 CTGATCTCTCTCGTCTCTTGTGCGGTGCGATCGTCTCCGAGCGCTCCCCAGCGCG 129

DB 83 CTGATCTCTCTCGTCTCTTGTGCGNGTGCATCTCTCCGAGCGCTCCCCAGCGCG 142

QY 130 CGCGCCCTCCGCGCGCTCGGCTCCGCGGCCGCTCTGCG-ACCAGCTGCTCTGCG 188

DB 143 CGCGCCCTCCGCGCGCTCGGCTCCGCGGCCGCTCTGCGAACACAGCTGCTCTGCG 202

Qy	189	CACTCGGCTACCCCTGCAACCGAGCACACGTTTGAAACAAGAATGGATTCTTTTATCTC	248
Db	203	CACTCGGGTACCCTGCACCGAGCACACGTTTGAAACAAGAATGGATTCTTTTATCTC	262
Qy	249	TTCAGCATATCCCACAT--GGCAAAAATAAAGCAGCAGATAGTACTGGGCCCTCCAGTTTTT	307
Db	263	TTCAGCATATCCCACATCGGCAAAAATAAAGCAGCAGATAGTACTGGGCCCTCCAGTTTTT	322
Qy	308	CTTCAACATGTCCTTTTTCAGGGAGGAGACACATGGTTTCATAAACTCTGCTGAGCAATCA	367
Db	323	CTTCAACATGTCCTTTTTCAGGGAGGAGACACATGGTTTCATAAACTCTGCTGNGCATCAC	382
Qy	368	CTTGGGTAT-ATCCCTGCTGATACGGTTTTCATCTTCTGTTTCATGA	465
Db	383	TTGGGGTATAATCTTCTGATACGGTTTTCATCTTCTGTTTCATGA	441
Qy	427	GCGTTGGAGTAAGGTCATTCAACCTTTCTCTGTTTCATGA	465
Db	442	GCGTTGGA--TAAGGTATCAACCTTTTCTAGTTTCATGA	478
RESULT 6			
BM037885		631 bp mRNA linear EST 06-NOV-2001	
LOCUS	Sll3f1l	Stem library from Oryza sativa (3-5 leaf stage)	Oryza
DEFINITION	sativa cDNA clone Sll3f1l, mRNA sequence.		
ACCESSION	BM037885		
VERSION	BM037885.1	GI:16753506	
KEYWORDS	EST.		
SOURCE	Oryza sativa.		
ORGANISM	Oryza sativa		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.		
AUTHORS	1 (bases 1 to 631) Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.		
TITLE	A Gene Expression Screen in Oryza sativa		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Dong HT Laboratory of Functional Genetics Bio-Technology Institute of Zhejiang University Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China Tel: 0086-571-86892051 Fax: 0086-571-86961525 Email: htdong@juem.zju.edu.cn Seq primer: M13 forward primer. Location/Qualifiers		
FEATURES	source	1..631	
		/organism="Oryza sativa"	
		/db_xref="taxon:4530"	
		/clone="Sll3f1l"	
		/clone_lib="Stem library from Oryza sativa (3-5 leaf stage)"	
		/tissue_type="Stem"	
		/dev_stage="3-5 leaf stage"	
		/note="Vector: pSport2"	
BASE COUNT	177 a	117 c	153 g
ORIGIN		174 t	10 others
Query Match	23.0%;	Score 341.6;	DB 13; Length 631;
Best Local Similarity	84.7%;	Pred No. 3.3e-69;	
Matches	466;	Conservative 0;	Mismatches 65; Indels 19; Caps 7;
Qy	820	CAACAATTGCTATCTCGGATTACAGGGGAAAACTGGTTCTCAATACATCAAGGATTGA	879
Db	73	CACCAAAATGTTNTATCGCGGATACGGGGAAACTGGGCTTCAATACATCAAGG-----	127
Qy	880	TTATTATTGGAGTATGAACCTCATCATCATCA-CAAAAAATCTGCACCATCTTTTC	938
Db	128	--GTTATTATGGGAGTNGGACTCATGATAGTAGTCGACGAAAAATTTGCACCATCGTTTTT	185

XhoI; Immature seeds, collected from pods ranging in age from 25 to 35 days after pollination, were harvested from greenhouse-grown plants. Seed were removed and separated from pod walls and immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

```
BASE COUNT      228 a 170 c 172 g 258 t
ORIGIN
Query Match      22.1%; Score 327.6; DB 14; Length 828;
Best Local Similarity 68.4%; Pred. No. 5.2e-66;
Matches 469; Conservative 0; Mismatches 214; Indels 3; Gaps 1;

Qy 293 GGCCTCCAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGACACATGGTTTCATAAAC 352
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 GGTCTCCAGTTCTGCTTCTCCATGGCTCTTCATGCGCAGGTGATGCTGTTTCTAAT 170

Qy 353 TCTGCTGACATCACTTGGGTATATCCTTGTGTATACAGGTTTGTGATTTGGATTGG 412
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 ACACAGAACACTCATGTGGCTTCATCCTTGCAGATCAAGGTTTGTATCTTTGGGTAGGA 230

Qy 413 AATGTCGGTGGAAACGGTTGGAGTAAAGCTTATCAACCTTTTCTGTTTCATGATAAGCTT 472
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 AATGTCGGTGGAAACGGTTGGAGCCTATGGCATAATCTTATCAGAGAAGATTAAGCAT 290

Qy 473 TTCTGGGATGGAGTGGCAAGAGTTAGCTGAATATGACCTTTTATGCAATGCTAGGCTAT 532
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 TTTTGGGATGGAGTGGCAAGAAATAGCCCTGTATGATCTTGCAGAAATGATCAATAT 350

Qy 533 GTGTATACAGTACACAGTCCAAATTTCTATATGTGGGCGATTCACAGGAACTATATG 592
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 ATTAATTCAGTGACAACTCAAAGTTATTTGTAGTTGGGCATTCACAGGGGACAAATTATA 410

Qy 593 GGTGTTGGCGCTTTCAGCATGCCGAATAGTAAATAGTATTAGCTCTGCGACACTTCTT 652
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 TCTTTTGTGCTTCTACCAACCGGATATAGTAAAGGTTGAAGCTGCGACTTCTTG 470

Qy 653 TGTCTATTTCTTATCTTGTATCATGTTAGTGTAGTCTTTTGTCTCAGAGCATGCCCAT 712
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 TCTCCAATTTCACTATGATCATATAGTGCACCTTTGGTACTAAGAATGAAAGATG 530

Qy 713 CATCTTGATCAGATGCTGTTACTATGGAATTCACAGCTGAACCTTCGATAGGCACATG 772
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 CACATTGATCAGATGATTTCTACAATGGGTGTTTCATCAGCTGAATTTCAAAGCGGAT 590

Qy 773 GGGGTTCAAATAGTAGATTCTTTGTGGCATGGTGAACACGTCGATTTGCAACAATTTGCTA 832
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 GAAACCAATCTCTGATTTCTTATGTGATCT---CGTCTAAGTTGCGATGACATGCTG 647

Qy 833 TCTGGCATTTACAGGGGAAACACTGTTGCTTCAATACATCAAGGATTTGATTTATTTGGAG 892
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 TCATCCATTTACAGGAAAGAATTTGTTGCTTCAACGAGTCAACGCGTGGCTTATTATCTTGA 707

Qy 893 TATGACCTCATCCATCATCGACAAATAATTCGACCATCTTTTTCAGATGATCAGGAA 952
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 708 CAAGAACCTTCACCCCATATCCCTCCAAAACCTTGAACACCTTTTCCAGATGATCCGCAA 767

Qy 953 GGCACCTTTCGCAAGATGATGACTATGG 978
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 768 GGCACCTTCTCCAGTACGATATGG 793
```

```
RESULT 8
BM501471
LOCUS BM501471.1 441 bp mRNA linear EST 14-FEB-2002
DEFINITION PAC00000000563 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
ACCESSION BM501471
VERSION BM501471.1 GI:18661576
```

```
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 441)
AUTHORS Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and
Jung,R.
TITLE Maize opaque endosperm mutations create extensive changes in
patterns of gene expression
JOURNAL Unpublished (2002)
COMMENT Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.
FEATURES
source 1..441
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/note="vector: pSport1; Site_1: Sali; Site_2: NotI"
BASE COUNT 104 a 90 c 102 g 144 t 1 others
ORIGIN
Query Match 22.1%; Score 327.4; DB 13; Length 441;
Best Local Similarity 85.6%; Pred. No. 6.8e-66;
Matches 375; Conservative 0; Mismatches 62; Indels 1; Gaps 1;

Qy 290 ACTGGCCCTCCAGTTTCTTCAACATGGTCTTTTTCAGGGAGGAGACACATGGTTTCATA 349
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 ACGCGTCCGCCAGTTTCTTTCAGCAGCGTCTTTTCCAGGGTGGAGATACATGGTTTCATA 64

Qy 350 AACTCTGTGAGCAATCACTTGGGTATATCCTTGTGCTGATAACGGTTTGTGATTTGGATT 409
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 AACTCCAATGAACAATCACTTGGGATATATCCTTGTGCAATGTTTGTGATTTGGGTC 124

Qy 410 GGAATGTCCTGGACCGCTTGGAGTAAAGTCATCTCAACCTTTTCTGTTTCATGATTAAG 469
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 GGAATGTCCTGGACACAGTTGGAGTAAAGCCACTTACTCTCTCTGTTTCATGATTAAG 184

Qy 470 CTTTCTGGGATTTGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTATGCAATGCTAGGC 529
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 CTTTCTGGGATTTGAGCTGGCAAGACCTTGTGAATACGACGTTTGGCAATGTTAAGC 244

Qy 530 TATGTTATACAGTCACACAGTCCAAAATTTCTATATGTGGGCAATTCACAGGAACTATA 589
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 TATGTTATACAGTTGCACAGTCCAAAATTTTGTATGTGGGCAATTCACAGGAACTATC 304

Qy 590 ATGGTTTGGCGGCTTTGACGATGCCGAAATAGTAAATAATGATCTCTGCGACACTT 649
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 305 ATGGTTTGGCGGCTTTTACAATGCCTGAAACAGTAAAGATGATAAGCTCTGCGGCTT 364

Qy 650 CTTTGTCTCTATTTCTTATCTTGTATCATCTAGTCTGTTTGTCTCAGAGCATGCC 709
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 CTTTGTCCCAATTTCTTACCTTGTATCAGTCAGTCAG-NCTAGTTTGTCTTAGACAGTGGCC 423

Qy 710 ATGCATCTTGTATCAGATG 727
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 ATGCATCTTGTATGATG 441

RESULT 9
AUI161807
LOCUS AUI161807 365 bp mRNA linear EST 03-APR-2002
DEFINITION AUI161807 Rice shoot Oryza sativa (japonica cultivar-group) cDNA
clone S5674, mRNA sequence.
ACCESSION AUI161807
VERSION AUI161807.1 GI:10931533
KEYWORDS EST.
```



```

QY 377 ATCTTGTCTGATACAGGGTTTTGATGTTTGGATTGGAAATGCCGTGGAAACGCTTGGAGT 436
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 406 ATCTTGTCTGATAATGGTTTGTGATGTTTGGATTGGAATGCCGTGGCACACGTTGGAGT 465
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 437 AAAGTGATTAACCTTTTCTGTTTCATGATAGCTTTTCTGGGATGGAGCTGGCAAGAG 496
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 AAAGGCCACTCTACTCTCTCTGTTTCATGATAAGCTTTTCTGGGATGGAGTTGGCAAGAC 525
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 497 TTAGCTGAATAGACCTTTTAGCAATGCTAGGCTATGTTATACAGT 543
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 CTGCTGAATACGATGTTTGGCAATGTTAAGCTATGATATACAAAT 572
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
LOCUS BE367231 476 bp mRNA linear EST 20-JUL-2000
DEFINITION P11_43_D12.g2.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE367231
VERSION BE367231.1 GI:9308788
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 476)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
,L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTMix
High quality sequence start: 13
High quality sequence stop: 473
POLYA-No.
FEATURES
source
1..476
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Pathogen induced 1 (P11)"
/note="Organ: Anthracnose-infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;
Site_2: EcoRI; Two-week-old sorghum plants (BTX 623
cultivar) were infected with pathogen (isolate FRM421 of
Colletotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings 48 hours after inoculation. Note: young
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda Zap II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen."
BASE COUNT 133 a 97 c 115 g 131 t
ORIGIN
Query Match 21.3%; Score 316.4; DB 10; Length 476;
Best Local Similarity 83.4%; Pred. No. 2.5e-63;
Matches 372; Conservative 0; Mismatches 71; Indels 3; Gaps 1;

```

```

QY 826 TTTGCTATCTCGGATTACAGGGGAAACCTGTTGCTTCAATACATCAAGGATTGATTATTA 885
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TCTGTTATCTTCATATACAGGTCAAAACTGTTGTTTCAATTCATCTCGGATTGACTATTA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 886 TTTGGAGTATGAACCTTCATCCATCATCGACAAAAAATCTGCACCATCTTTTTCAGATGAT 945
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TTTGGAGTATGAACCCCATCCGTCATCAACAAAAAATTTGCGGCATCTTTTCCAGATGAT 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 946 CAGAAAGGCACCTTTCGCAAGTATGACTATGGTTATTGGGAAACCTAAGGCCCTACGG 1005
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CAGAAAGGCAGTTCGCAAGTATGACTATGGTGTGGGAAACATAAGGCACATACGG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1006 TCATTTGCTCCTCCGCAATTTGACCTAAGCAGCATACCGAAATCACTGCCCATATGGAT 1065
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CCAGCGCATCTCCCTCATTCGATCTAAGCAGCATACCGGAATCACTGCCAATATGGAT 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1066 GGGATATGGAGGCTTGTGATGCAATGGGTAAACCGATGTTTCAGCGTACTATCAGAGA 1125
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GGGTATGGAGGCTTGTGATGCACTGGCTGACGTAAACAGTCTTGAGCGCACCATCAAGA 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1126 GCTGGATCTACACGAACTTCTGTACATGCTGACTATGGCCATATGATTTTGTAT 1185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GTTGAGATCCAGCGCAGAGCTGCTGACATGGTGACTATGGCCACATTTGATTTTCATCAT 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1186 GAGCGTGAAGCGGAAAGATGATGTTTATGTGACCTAATAAGATTTCTTAGGG---AAAA 1242
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GAGCGTGAAGCGGAGGACGATGTTTATGTGACTTAATGAGTTTCTCAGGCGCCACCA 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1243 TGGATGCGATAATAGCTATTAGGATG 1268
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GGAATGCACAGTAGTTATTAGGACG 446
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
LOCUS BM328649 647 bp mRNA linear EST 04-JAN-2002
DEFINITION P1C1_24_G11-gl.A002 Pathogen-infected compatible 1 (PIC1) Sorghum
bicolor cDNA, mRNA sequence.
ACCESSION BM328649
VERSION BM328649.1 GI:18067786
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 647)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
Sudman,M. and Pratt,L.H.
TITLE An EST database from Sorghum: plants infected with a compatible
pathogen
JOURNAL Unpublished (2002)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector, and regions
below Phred quality 16. The threshold for highest quality sequence
is 20. Three-prime sequences, which are obtained with PolyTMix or
T7 sequencing primer, are presented as the reverse complement.
Seq primer: 17
High quality sequence start: 61
High quality sequence stop: 647
POLYA=Yes.
FEATURES
source
1..647
/organism="Sorghum bicolor"
/cultivar="Btx623"
/db_xref="taxon:4558"

```


/clone_lib="Pathogen-infected compatible 1 (PIC1)"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with
Colletotrichum graminicola"
/note="vector: pBluescript II SK(-) from Lambda Zap II;
Site_1: XhoI; Site_2: EcoRI; Four-week-old sorghum
seedlings were sprayed with spore suspension prepared from
3-week-old FRM421, a sorghum isolate of the anthracnose
pathogen Colletotrichum graminicola. Inoculated plants
were kept in a 25 C dark growth chamber with 100% relative
humidity for 24 hr, followed by 12/12 hr of light/dark
cycle at 25 C with 90% relative humidity for another 24
hr. All leaves were harvested and quick frozen with liquid
nitrogen and stored in a -80 C freezer. The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision.
WARNING: While most or all ESTs are expected to derive
from the host plant, no effort was made to eliminate ESTs
deriving from the pathogen."
BASE COUNT 182 a 128 c 158 g 179 t
ORIGIN

Query Match 19.9%; Score 294.4; DB 13; Length 647;
Best Local Similarity 78.7%; Pred. No. 3.5e-58;
Matches 365; Conservative 0; Mismatches 96; Indels 3; Gaps 1;
Qy 873 GGATTGATTATTATTGGAGTATGAACCTCATCCATCATCGACAAAATCTGCACCATC 932
Db 2 GGATTGACTATTATTGGAGTATGAACCCCATCGTCATCAACAAAATTTGGCGCATC 61
Qy 933 TTTTTCAGATGATCAGGAAGGCACTTTTCGCAAAATGACTATGGTATTGGGAAC 992
Db 62 TTTTCCAGATGATCAGGAAGGCACTTTTCGCAAAATGACTATGGTATTGGGAACA 121
Qy 993 TAAGCGCTACGGTCATTTCGCTCCTCCCGCATTTGACCTTAAGCAGCATACCAATCAC 1052
Db 122 TAAGCACTACGGCCAGGGCATCTCCCTCATTCGATCTAAGCAGCATACCGGAATCAC 181
Qy 1053 TGCCCATATGGATGGATATGAGGCTTTGATGCTTGGCTGATGTAACCGATGTTACG 1112
Db 182 TGCCAATATGGATGGGATATGAGGCTTTGATGCTTGGCTGATGTAACCGATGTTGAGC 241
Qy 1113 GTACTATCAGAGCTGGGATCTACACCAAGACTTCTGTACATGGTGACTATGGCCATA 1172
Db 242 TCACCATCAAAAGTTGAGATCCACGCCAGAGCTGCTGTACATGGTGACTATGGCCACA 301
Qy 1173 TTGATTTTGTTATGAGCTGAAGCGGAAGATGATGTTTATGTGACCTTAATACATTC 1232
Db 302 TTGATTTCATCAGCGTGAAGCGGAAGGACGATGTTTATGTGACCTTAATAGGTTTC 361
Qy 1233 TTAGG---AAAATGGATGGCATATAGCTATTAGGATGCTTTCATGTGTATAATAAAA 1289
Db 362 TCAGGGCCACAGGAATGCACAGTAGTTATTAGGAGCATCGTATGCCCTTAATTAA 421
Qy 1290 CATCTGTACAGTATTGGTCTCTCCCGATGAGTATGTATATA 1333
Db 422 TGTATTACTATGATGACGATACGATGATGTAAATACGTACATA 465

RESULT 13
AU161802
LOCUS AU161802 323 bp mRNA linear EST 03-APR-2002
DEFINITION AU161802 Rice shoot Oryza sativa (japonica cultivar-group) cDNA
clone S5639, mRNA sequence.
ACCESSION AU161802
VERSION AU161802.1 GI:10931528
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group).
ORGANISM Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 323)

AUTHORS Sasaki, T. and Yamamoto, K.
TITLE Rice cDNA from etiolated shoot (2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@nri.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT "RGP"
Location/Qualifiers
1. 323
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="S5639"
/clone_lib="Rice shoot"
/dev_stage="Etiolated shoot (8 days old)"
/note="Etiolated shoot (8 days old)"
BASE COUNT 59 a 119 c 83 g 61 t
ORIGIN 1 others

Query Match 19.3%; Score 285.8; DB 9; Length 323;
Best Local Similarity 99.0%; Pred. No. 3.5e-56;
Matches 298; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 10 CACAGCGCGGGCGGTGGCGATGGCGATGGCGGCCACGCCCGGAGGAGCGCTCCC 69
Db 23 CAAAGCGCGGGCGGTGGCGATGGCGATGGCGGCCACG-CCCCGGAGGAGCGCTCCC 81
Qy 70 CCTGATCCTCTCGTCTCTTGTCTGGGTGCGATCGTCTCGGAGGCTCCCCAGCGC 129
Db 82 CCTGATCCTCTCGTCTCTTGTCTGGGTGCGATCGTCTCGGAGGCTCCCCAGCGC 141
Qy 130 GCGCGCCCTCCGCGCGTCCGCTCCGCTCCGCGGCGCTCTCGACACGCTCTCTCT 189
Db 142 GCGCGCCCTCCGCGCGTCCGCTCCGCTCCGCGGCGCTCTCGACACGCTCTCTCT 201
Qy 190 ACTCGGCTACCCCTCGACCGAGCACAGTTGAACAAAGATGATTCCTTTATCTCT 249
Db 202 ACTCGGCTACCCCTCGACCGAGCACAGTTGAACAAAGATGATTCCTTTATCTCT 261
Qy 250 TCAGCATATCCACATGCGCAAAAATAAAGACGACAGATAGTACTGGCCCTCCAGTTTTTCT 309
Db 262 TCAGCATATCCACATGCGCAAAAATAAAGACGACAGATAGTACTGGCCCTCCAGTTTTT 321
Qy 310 T 310
Db 322 T 322

RESULT 14
BI075485
LOCUS BI075485 646 bp mRNA linear EST 20-JUN-2001
DEFINITION IP1_21_B03.g1_A002 Immature pannicle 1 (IPI) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BI075485
VERSION BI075485.1 GI:14514142
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 646)
AUTHORS Klein, R.R., Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M. and Pratt
L.H.
TITLE An EST database from Sorghum: developing preanthesis pannicles
JOURNAL Unpublished (2001)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement.

Seq primer: T7
High quality sequence start: 12
High quality sequence stop: 646
POLYA=Yes.

FEATURES	Location/Qualifiers
source	1. .646

```

1: .089
/organism="Sorghum bicolor"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone_lib="Immature pannicle 1 (IPI)"
/notes="Organ: Developing preanthesis pannicles; Vector:
pBluescript II SK(-) from Lambda zap II; Site_1: XhoI;
Site_2: EcoRI; The library was made from poly-A RNA in the
cloning vector lambda zap II. Clones to be sequenced were
prepared by mass excision."
182 a 132 c 156 g 176 t
BASE COUNT

```

BASE COUNT	182 a	132 c	156 g	176 t
ORIGIN				
Query Match	19.28;	Score 285.2;	DB 13;	Length 646;
Best Local Similarity	77.94;	Pred. No. 4.9e-36;		
Matches 357;	Conservative	0;	Mismatches 98;	Indels 3;
				Caps 1;

Qy	879	ATTATATTTGGAGTATGAACCTCATCCATCATCGACAAAATAATCTGCACCATCTTTTTC	938
Db	1	ACTATTTATTTGGAGTATGAACCCCATCGTATCAACAACAAAATAATTCGGCATCTTTTTC	60
Qy	939	AGATGATCAGAAAGGACATTTCCCAAGTATGACTATGGGTATATGGGAACCTAAGGC	998
Db	61	AGATGATCAGAAAGGACATTTCCCAAGTATGACTATGGATGGTGGGGAACCATAGGC	120
Qy	999	GCTACGGTCATTTTCGCTCTCCGCATTTGACCTTAAGCAGCATACCAAGATCATCTGCCCA	1058
Db	121	ACTAGGCCACGGCGATCCTCCCTCATTTCCGATCTAAGCAGCATACCGGAATCATCTGCCAA	180
Qy	1059	TATGGATGGGATATGGAGGCTCTTGATCATTTGGCTGATGAACCGATGTTTCAGCGCTACTA	1118
Db	181	TATGGATGGGATATGGAGGCTCTTGATGACCTGGCTGAGTAAACAGATGTTGAGCGCACCA	240
Qy	1119	TCAGAGAGCTGGGATCTACACAGAACTCTGTACATTTGGTGACTATGGCCATATTTGATT	1178
Db	241	TCAAGAGTTGAGATCCACGCCAGAGCTGCTGTACATTTGGTGACTATGGCCACATTTGATT	300
Qy	1179	TGTTATGAGCGGTGAAGCGGAAGATGATGTTATGTGGACCTTAATAGATTTCTTAGGG	1238
Db	301	TCATCATGAGCGGTGAAGCGGAAGCAGATGTTATGTTGACTCTATGAGGTTTCTCAGGG	360
Qy	1239	-- -AAAAATGGATGGCATAATAGCTATTAGGATGCTTCTCATGTGTATATAAAAAACATCTG	1295
Db	361	CCCAGCGGGAATGCACAGTAGTTATTATAGGACGATCGTATGCGCTTTAATTAATGTTTA	420
Qy	1296	TACAGTATTGGTCTCTCCCGATGTGAGTATGTATATA	1333
Db	421	TACTATGATGACAGTACGACTATGCTTAATACGTACATA	458

RESULT 15	AU088738	521 bp	linear	EST 02-APR-2002
LOCUS	AU088738			
DEFINITION	AU088738 Rice shoot Oryza sativa (japonica cultivar-group) cDNA			
	clone S5305, mRNA sequence.			
ACCESSION	AU088738			
VERSION	AU088738.1	GI:7378467		
KEYWORDS	EST.			
SOURCE	Oryza sativa (japonica cultivar-group).			
ORGANISM	Oryza sativa (japonica cultivar-group)			

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; *Oryza*.
1 (bases 1 to 521)
Sasaki, T. and Yamamoto, K.
Rice cDNA from etiolated shoot (2000)
Unpublished (2000)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@agr.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>
PROJECT = 'rgp'

FEATURES source

```
1 521
/organism="oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="S5303"
/clone_lib="Rice shoot"
/dev_stage="Etiolated shoot (8 days old)"
```

BASE COUNT	149 a	111 c	100 g	160 t	1 others
ORIGIN					

Query Match	18.9%;	Score	281;	DB	9;	Length	521;		
Best Local Similarity	90.7%;	Pred. No.	4.7e-55;						
Matches	332;	Conservative	0;	Mismatches	31;	Indels	3;	Gaps	3;

QY	174	ACCAGCTGCTCCTGCACCTCGGCTACCCCTGCACCGAGCACAAACGTTGAAAACAAAAGATG	233
Db	7	AACAGCTNGTCTCCAACTCGGTACCTGCAGGAG-TCACAACGTTGAAAACAAAAGATG	65
QY	234	GATTCCTTTTTATCTCTTCAGCATATCCCACTGGCAAAAATAAAGCAGACGATGACTGTG	293
Db	66	GATTCCTCTAACTCTTCCAGCATATTCCCACATATGCAAAAATAAAGCAGACGATGACT-	124
QY	294	GCCCTCAGTTTTTCTTCAACATGGCTTTTTCAGGGAGGACACATGGTTTCATAAACT	353
Db	125	GCCCTCCAGTTTTTCTTCAACAT-GTCTTTTTCAGGGAGGACACATGGTTTCATAAACT	183
QY	354	CTGCTGAGCAATCACTTGGGTATATCCCTTGCTGATACCGTTTTTGATGTTGGATGGGA	413
Db	184	CTGCTGAGCAATCACTTGGGTATATCCTTGCTGATACCGTTTTTGATGTTGGATGGGA	243
QY	414	ATGTCCTGGAAACGGTTGGAGTAAGGTCATTCACCTTTTCTGTTTCATGATAAGCTTT	473
Db	244	ATGTCCTGGAAACGGTTGGAGTAAGGTCATTCACCTTTTCTGTTTCATGATAAGCTTT	303
QY	474	TCTGGGATTGGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTAGCAATGCTTAGGCTATG	533
Db	304	TCTGGGATTGGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTAGCAATGCTTAGGCTATG	363
QY	534	TGTATA	539
Db	364	TATATA	369

Search completed: December 30, 2002, 19:04:44
Job time : 2235 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2002, 17:14:21 ; Search time 77 seconds
(without alignments)
5906.515 Million cell updates/sec

Title: US-09-699-652a-13
Perfect score: 1483
Sequence: 1 gcacgagtcacagcgcgc.....ctctcaaaaaaaaaaaaaa 1483
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91.2	6.1	1134	US-09-186-489-1	Sequence 1, Appli
2	91.2	6.1	1411	US-09-186-489-3	Sequence 3, Appli
3	91.2	6.1	8324	US-09-186-489-5	Sequence 5, Appli
4	88	5.9	1134	US-09-186-489-4	Sequence 4, Appli
5	81.6	5.5	1360	US-09-820-001-1	Sequence 1, Appli
6	64.4	4.3	1137	US-08-227-108-2	Sequence 2, Appli
7	64.4	4.3	1137	US-09-073-674-2	Sequence 2, Appli
8	64.4	4.3	1140	US-08-227-108-4	Sequence 4, Appli
9	64.4	4.3	1146	US-09-073-674-4	Sequence 4, Appli
10	64.4	4.3	1146	US-08-227-108-6	Sequence 6, Appli
11	64.4	4.3	1146	US-09-073-674-6	Sequence 6, Appli
12	64.4	4.3	1528	US-08-227-108-1	Sequence 1, Appli
13	64.4	4.3	1528	US-09-073-674-1	Sequence 1, Appli
14	52.2	3.5	22067	US-09-820-001-3	Sequence 3, Appli
15	46.4	3.1	1926	US-09-249-585A-4	Sequence 4, Appli
16	46.4	3.1	1931	US-09-130-114-2	Sequence 2, Appli
17	45	3.0	2914	US-09-177-650-6	Sequence 6, Appli
18	44.8	3.0	7218	US-08-232-463-14	Sequence 14, Appli
19	43	2.9	1112	US-09-173-300-48	Sequence 48, Appli
20	43	2.9	1333	US-09-372-422A-9	Sequence 9, Appli
21	42.6	2.9	1439	US-09-056-556-167	Sequence 167, App
22	42.6	2.9	1439	US-09-072-596-162	Sequence 162, App
23	42	2.8	2220	US-08-765-907A-14	Sequence 14, Appli
24	42	2.8	2888	US-08-765-907A-1	Sequence 1, Appli
25	42	2.8	15872	US-09-105-537-1	Sequence 1, Appli
26	42	2.8	28958	US-08-258-261B-6	Sequence 6, Appli
27	42	2.8	28958	US-08-456-837-6	Sequence 6, Appli

28	42	2.8	28958	1	US-08-457-342-6	Sequence 6, Appli
29	42	2.8	28958	1	US-08-457-646A-6	Sequence 6, Appli
30	42	2.8	28958	1	US-08-458-076A-6	Sequence 6, Appli
31	42	2.8	28958	1	US-08-764-233A-4	Sequence 4, Appli
32	42	2.8	28958	1	US-08-457-335A-6	Sequence 6, Appli
33	42	2.8	28958	1	US-08-729-214-6	Sequence 6, Appli
34	42	2.8	28958	3	US-09-028-934-6	Sequence 6, Appli
35	42	2.8	49377	1	US-08-764-233A-1	Sequence 1, Appli
36	41.8	2.8	1417	4	US-09-199-737-3	Sequence 3, Appli
37	41.8	2.8	1417	4	US-09-058-333A-3	Sequence 3, Appli
38	41.4	2.8	668	6	5498529-5	Patent No. 5498529
39	41.2	2.8	4362	2	US-08-455-073A-1	Sequence 1, Appli
c 40	41	2.8	11219	1	US-07-642-734C-1	Sequence 1, Appli
c 41	41	2.8	11219	3	US-08-439-009A-1	Sequence 1, Appli
c 42	40.8	2.8	907	1	US-08-821-119-1	Sequence 1, Appli
c 43	40.6	2.7	614	4	US-08-998-416-861	Sequence 861, App
c 44	40.6	2.7	851	4	US-08-818-112-33	Sequence 33, Appl
c 45	40.6	2.7	851	4	US-08-818-111-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-186-489-1
; Sequence 1, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186.489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Kid (Goat)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1134)
US-09-186-489-1

Query Match	6.1%	Score 91.2;	DB 4;	Length 1134;
Best Local Similarity	54.2%	Pred. NO. 1.9e-14;		
Matches 230;	Conservative	0;	Mismatches 188;	Indels 6;
Gaps	2;			
OY	194	GGCTACCCCTGCACCGACACACCGTTGAAACAAAGATGGATTCCTTTATCTCTTCAG	253	
Db	67	GGCTACCCCAAGTGAGATCGATAAAGTTATTAAGTCAGATGGCTATATCTTCAGGTCTAT	126	
OY	254	CATATCCACATGGCAAAATAAAGCAGCAGAT---AGTACTGGCCCTCCAGTTTCTT	310	
Db	127	CGGATTCCTCATGGAAGAAGATGATGCTATCATTTAGGTCAGAGACCTTCTGTTCTG	186	
OY	311	CAACATGGTCTTTTTCAGGGAGGAGACACATGGTTCATAAAGCTGTGTCAGCAATCACTT	370	
Db	187	CAGCATGGTCTTCTTTCCTCAGCTACAACTGGATTTCCAACTTCCCAACACACGCTG	246	
OY	371	GGGTATATCTTGCTGATAACGGTTTGTGATTTGGATTGGCAATGTCGTCGAACGGT	430	
Db	247	GGCTTCCTCCCTGGCAGATGCTGGTTATGAGCTGGTGGGGAACAGCAGAGAAACACT	306	
OY	431	TGGAGTAAAGGTCATTCAACCTTTTCTGTCATGATAAGCTTTTCTGGGATTGGAGCTGG	490	
Db	307	TGGGCCCAGGACATTTATACCTATTACCAGACTCCCTGGAATTCGGCTTTCAGCTTT	366	
OY	491	CAAGAGTAGCTGAATATGACCTTTTAGCAATGCTAGGCTATGTGT---ATACGTGACA	547	

Db 367 GATGAATGGCTGAATATGACGCTTCCATCTACAATTGATTTTCATCTTAAGAGAACAGGA 426
Qy 548 CAGTCCAAAATCTATATCTGGGGCATTCACAGGGAACATAATGGGTTTGGCGGCTTTG 607
Db 427 CAGAAGAAGCTACACTATGTTGGCCATTCCCAAGGACACCAATGGTTTGTGCGCCTT 486
Qy 608 ACGA 611
Db 487 TCTA 490

RESULT 2

US-09-186-489-3

; Sequence 3, Application US/09186489

; Patent No. 6375947

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul L.

; APPLICANT: Scharpf Jr., Lewis G.

; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and

; TITLE OF INVENTION: Processes for its Production and Use

; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186.489

; CURRENT FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1411

; TYPE: DNA

; ORGANISM: Kid (Goat)

US-09-186-489-3

Query Match 6.1%; Score 91.2; DB 4; Length 1411;
Best Local Similarity 54.2%; Pred. No. 2e-14;
Matches 230; Conservative 0; Mismatches 188; Indels 6; Gaps 2;

Qy 194 GGCTACCCCTGCACCGAGACACACGTTGAAACAAAGATGGATTCCTTTATCTCTTCAG 253
Db 191 GGCTACCCCAAGTGAGATGCATAAAGTTATACTGCAGATGCTATATCTTCAGGCTCTAT 250
Qy 254 CATATCCACATGGCAAAATAAAGCAGCAGAT---AGTACTGGCCCTCCAGTTTCTT 310
Db 251 CGGATTCCTCATGGAAGAATGATGCTAATCATTTAGGTCAGAGACCTGTGTGTTCTG 310
Qy 311 CAACATGCTCTTTTCAGGGAGGAGACACATGGTTTCATAAAGCTGCTGAGCAATCACTT 370
Db 311 CAGCATGCTCTCTTCGCTCAGCTACAACTGGATTTCCACCTTCCCAACAACAGCCTG 370
Qy 371 GGGTATATCCTTGTGATAACGGTTTGTGATTTGGATTGGGAATGTCCTGGGAACGCGT 430
Db 371 GGCTTCCCTGGCAGATGCTGTTTATGACGTGTGCTGGGGAACAGCAGAGGAACACT 430
Qy 431 TGGAGTAAAGTCAATCAACCTTTCTGTTTCATGATAAGCTTTTCTGGGATTTGGAGCTGG 490
Db 431 TGGGCCCAAGGAACATTTATCTATTTACCTATTTACACGACTCCCTGAAATCTGGGCTTT 490
Qy 491 CAAGAGTTAGCTGAATATGACCTTTTAGCAATGCTAGGCTATGCT---ATACAGTCACA 547
Db 491 GATGAATGGCTGAATATGACCTTCCATCTCAATTTGATTTCAATCTTAAGAGAACAGGA 550
Qy 548 CAGTCCAAAATCTATATGTGGGCAATTCACAGGGAACATAATGGGTTTGGCGGCTTTG 607
Db 551 CAGAAGAAGCTACACTATGTTGGCCATTCCCAAGGACCACTTGGTTTGTGCGCCTTT 610
Qy 608 ACGA 611
Db 611 TCTA 614

RESULT 3

US-09-186-489-5

; Sequence 5, Application US/09186489

; Patent No. 6375947

; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186.489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 8324
; TYPE: DNA
; ORGANISM: Yeast YE-1 expression vector
US-09-186-489-5

Query Match 6.1%; Score 91.2; DB 4; Length 8324;
Best Local Similarity 54.2%; Pred. No. 4.4e-14;
Matches 230; Conservative 0; Mismatches 188; Indels 6; Gaps 2;

Qy 194 GGCTACCCCTGCACCGAGACACACGTTGAAACAAAGATGGATTCCTTTATCTCTTCAG 253
Db 1598 GGCTACCCCAAGTGAGATGCATAAAGTTATACTGCAGATGCTATATCTTCAGGCTCTAT 1657
Qy 254 CATATCCACATGGCAAAATAAAGCAGCAGAT---AGTACTGGCCCTCCAGTTTCTTCTT 310
Db 1658 CGGATTCCTCATGGAAGAATGATGCTAATCATTTAGGTCAGAGACCTGTGTGTTCTG 1717
Qy 311 CAACATGCTCTTTTCAGGGAGGAGACACATGGTTTCATAAAGCTGCTGAGCAATCACTT 370
Db 1718 CAGCATGCTCTCTTCGCTCAGCTACAACTGGATTTCCACCTTCCCAACAACAGCCTG 1777
Qy 371 GGGTATATCCTTGTGATAACGGTTTGTGATTTGGATTGGGAATGTCCTGGGAACGCGT 430
Db 1778 GGCTTCCCTGGCAGATGCTGTTTATGACGTGTGCTGGGGAACAGCAGAGGAACACT 1837
Qy 431 TGGAGTAAAGCTCAATCAACCTTTCTGTTTCATGATAAGCTTTTCTGGGATTTGGAGCTGG 490
Db 1838 TGGGCCCAAGGAACATTTATCTATTTACGACTCCCTGAAATCTGGGCTTTTCAGCTTT 1897
Qy 491 CAAGAGTTAGCTGAATATGACCTTTTAGCAATGCTAGGCTATGCT---ATACAGTCACA 547
Db 1898 GATGAATGGCTGAATATGACCTTCCATCTCAATTTGATTTCAATCTTAAGAGAACAGGA 1957
Qy 548 CAGTCCAAAATCTATATGTGGGCAATTCACAGGGAACATAATGGGTTTGGCGGCTTTG 607
Db 1958 CAGAAGAAGCTACACTATGTTGGCCATTCCCAAGGACCACTTGGTTTGTGCGCCTTT 2017
Qy 608 ACGA 611
Db 2018 TCTA 2021

RESULT 4

US-09-186-489-4

; Sequence 4, Application US/09186489

; Patent No. 6375947

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul L.

; APPLICANT: Scharpf Jr., Lewis G.

; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and

; TITLE OF INVENTION: Processes for its Production and Use

; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186.489

; CURRENT FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 1134

; TYPE: DNA

; ORGANISM: Bovine

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-08-227-108-4

Query Match 4.3%; Score 64.4; DB 1; Length 1140;
Best Local Similarity 49.8%; Pred. No. 1.9e-07;
Matches 220; Conservative 0; Mismatches 216; Indels 6; Gaps 2;
QY 176 CAGCTGCTCCTGCCACTCGGCTACCCCTGCACCGAGACACAGCTTGAACAAAGATGGA 235
DB 55 CAGATGATCACCTACTCTGGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACGGT 114
QY 236 TTCCCTTTTATCTCTCAGCATATCCACATGCGCAAAAATAAAGCAGCAGATA---GTACT 292
DB 115 TATATCTTGGGATCGACAGAATTCCTTATGGGAGGAAAAATTCAGAGAAATATAGGCCGG 174
QY 293 GGCCCTCCAGTTTTTCTTCAACATGCTCTTTTTCAGGGAGGAGACACATGTTTCATAAAC 352
DB 175 AGACCTGTGCAATTTTTCACACCGGTTTTCGCGATCGACCAAACTGGATCTCCCAAC 234
QY 353 TCTGCTGACCAATCACTTTGGGTATATCCTTGTCTGATAACGGTTTTCATGTTTGGATTGGG 412
DB 235 CTGCCCAACACAGCTGSCCTTCATCTCTGCGCGACGCGGTACGACGTGTGGCTGGGG 294
QY 413 AATGTCGTTGGAACGGTTGGAGTAAGGTCAATCAACCTTTTCTGTTTCATGATAGCTTT 472
DB 295 AACAGCAGGGGCAACACCTGGGCCAGGAGGAATCTGTACTCTGCCCGACTCCGTCGAA 354
QY 473 TTCTGGATTGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTAGCATGTAGCT-- 530
DB 355 TTCTGGCTTTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATTGACTTC 414
QY 531 -ATGTGTATACAGTACACAGTCCAAAATTCATATGTTGGGCAATTCACAGGAACTATA 589
DB 415 ATCTTGAAGAAAACGGACAGCAAGCTACACTAGCTTGGCCATTCACAGGCAACACC 474
QY 590 ATGGGTTTGGCGGCTTTGACGA 611
DB 475 ATTGGTTTCATCGCCTTTTCCA 496

RESULT 9
US-09-073-674-4
Sequence 4, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-09-073-674-4
Query Match 4.3%; Score 64.4; DB 2; Length 1140;
Best Local Similarity 49.8%; Pred. No. 1.9e-07;
Matches 220; Conservative 0; Mismatches 216; Indels 6; Gaps 2;
QY 176 CAGCTGCTCCTGCCACTCGGCTACCCCTGCACCGAGACACAGCTTGAACAAAGATGGA 235
DB 55 CAGATGATCACCTACTCTGGGGATACCCAGCTGAGGAATATGAAGTTGTGACCGAAGACGGT 114
QY 236 TTCCCTTTTATCTCTCAGCATATCCACATGCGCAAAAATAAAGCAGCAGATA---GTACT 292
DB 115 TATATCTTGGGATCGACAGAATTCCTTATGGGAGGAAAAATTCAGAGAAATATAGGCCGG 174
QY 293 GGCCCTCCAGTTTTTCTTCAACATGCTCTTTTTCAGGGAGGAGACACATGTTTCATAAAC 352
DB 175 AGACCTGTGCAATTTTTCACACCGGTTTTCGCGATCGACCAAACTGGATCTCCCAAC 234
QY 353 TCTGCTGACCAATCACTTTGGGTATATCCTTGTCTGATAACGGTTTTCATGTTTGGATTGGG 412
DB 235 CTGCCCAACACAGCTGSCCTTCATCTCTGCGCGACGCGGTACGACGTGTGGCTGGGG 294
QY 413 AATGTCGTTGGAACGGTTGGAGTAAGGTCAATCAACCTTTTCTGTTTCATGATAGCTTT 472
DB 295 AACAGCAGGGGCAACACCTGGGCCAGGAGGAATCTGTACTCTGCCCGACTCCGTCGAA 354
QY 473 TTCTGGATTGAGCTGGCAAGAGTTAGCTGAATATGACCTTTTAGCATGTAGCT-- 530
DB 355 TTCTGGCTTTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCATTGACTTC 414
QY 531 -ATGTGTATACAGTACACAGTCCAAAATTCATATGTTGGGCAATTCACAGGAACTATA 589
DB 415 ATCTTGAAGAAAACGGACAGCAAGCTACACTAGCTTGGCCATTCACAGGCAACACC 474
QY 590 ATGGGTTTGGCGGCTTTGACGA 611
DB 475 ATTGGTTTCATCGCCTTTTCCA 496
RESULT 10
US-08-227-108-6
Sequence 6, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21

Db	352	TTCTGGCGCTTTCAGCTTTGACGAGATGGCTAAATATGACCTTCCCGCCACCAATTGACTTC	411
Qy	531	-ATGTGTATACAGTCACAGATCCAAATTTATATGTGGGCATTCACAGGGAACATATA	589
Db	412	ATCTTGAAGAAACGGACAGCAAGCTACACTACGTTGGCCATTCCTCCAGGCGACACC	471
Qy	590	ATGGGTTTGGCGGCTTTGACGA	611
Db	472	ATTGGTTTCATCGCCTTTTCCA	493
RESULT 13			
US-09-073-674-1			
; Sequence 1, Application US/09073674			
; Patent No. 5998189			
; GENERAL INFORMATION:			
; APPLICANT: Blanchard, Claire			
; APPLICANT: Benicourt, Claude			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Warner-Lambert Company			
; STREET: 2800 Plymouth Road			
; CITY: Ann Arbor			
; STATE: Michigan			
; COUNTRY: U.S.A.			
; ZIP: 48105			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/073,674			
; FILING DATE:			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Crissey, Todd M.			
; REGISTRATION NUMBER: 37,807			
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 734 622-7530			
; TELEFAX: 734 622-1553			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1528 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: cDNA to mRNA			
US-09-073-674-1			
Query Match			
Best Local Similarity 49.8%; DB 2; Length 1528;			
Matches 220; Conservative 0; Mismatches 216; Indels 6; Gaps			
Qy	176	CAGCTGCTCTCTGCGACTCGCGCTACCCCTGCACCGACGACACAGTTGAAACAAAGATGGA	235
Db	52	CAGATGATGACCTACTGGGGATATGAACTGTGTGACCGAAGACGGT	111
Qy	236	TTCTCTTTTATCTCTTCAGCATATCCACATGGCAAAAATAAAGCAGACAGATA--CTACT	292
Db	112	TATATCTTGGGATCGACAGAAATCTCTATGGGAGGAAAAATTCAGAGAAATATAGGCCGG	171
Qy	293	GGCCCTCCAGCTTTTCTTCAACATGTCCTTTTTTCAGGGAGGAGACACATGGTTTCATAAAC	352
Db	172	AGACCTGTCGCAATTTTGGCAACAGCGTTGCTCGCATGCCACACAACCTGGATCTCCAAC	231
Qy	353	TCGTGCTGAGCAATACATTTGGGTATATCCTTGCTGATAACGGTTTTGATGTTTGGATTTGG	412
Db	232	CTGCCCAACAAACAGCTTGGCTTCACTCCTGGCCAGCGCGGGTACGACGTTGGCTGGGG	291

